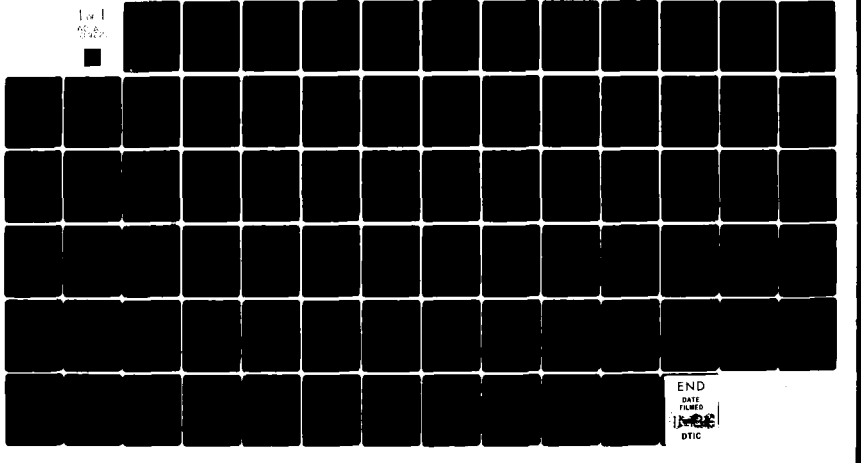
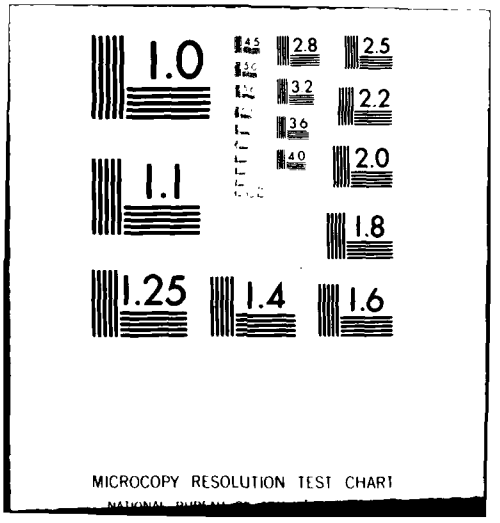
AD-A092 221 STANFORD UNIV CA DEPT OF OPERATIONS RESEARCH F/G 12/1
AN APPROACH TO REGENERATIVE SIMULATION ON A GENERAL STATE SPACE--ETC(U)
JUL 80 P W GLYNN N00014-76-C-0578
UNCLASSIFIED TR-53 NL



END
DATE FILMED
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART

AN APPROACH TO REGENERATIVE SIMULATION ON A GENERAL STATE SPACE

by

Peter W. Glynn

TECHNICAL REPORT NO. 53

July 1980

Prepared under Contract N00014-76-C-0578 (NR 042-343)

for the

Office of Naval Research

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

*This research was also partially supported under National Science Foundation Grant MCS79-09139 and a Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: TR-53	2. GOVT ACCESSION NO.: AD-A092221	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): AN APPROACH TO REGENERATIVE SIMULATION ON A GENERAL STATE SPACE	5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): PETER W. GLYNN	8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0578, NSF-MCS79-09139
9. PERFORMING ORGANIZATION NAME AND ADDRESS: DEPARTMENT OF OPERATIONS RESEARCH, STANFORD UNIVERSITY, STANFORD, CA 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 042-343
11. CONTROLLING OFFICE NAME AND ADDRESS: STATISTICS AND PROBABILITY PROGRAM, OFFICE OF NAVAL RESEARCH (Code 436), ARLINGTON, VA 20360	12. REPORT DATE: JUL 1980
	13. NUMBER OF PAGES: 17
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): 12 79	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Recurrent Markov Processes
Regenerative Method
Simulation
Statistical Estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

402766

A wide variety of stochastic systems may be viewed as Markov chains taking on values in a general state space. An example is the class of generalized semi-Markov processes, which are commonly obtained in network queueing problems via the technique of supplementary variables. A simulator is often interested in obtaining steady state properties of such a system. Some recent developments in Markov chain theory by Athreya, Ney, and Nummelin allow one to embed a certain subclass of these processes in a regenerative environment. We study some consequences of this embedding and develop statistical estimation procedures for the general problem that bear close resemblance to the regenerative method of simulation analysis for finite state Markov chains.

1. Introduction

Many complex stochastic systems of interest to simulators involve general state space considerations. For example, the general discrete-event simulation, when viewed as a generalized semi-Markov process, is basically a discrete time Markov chain living on an uncountable state space.

In this paper, we shall consider a class of general state space Markov chains satisfying a certain recurrence property known as Harris recurrence. We shall explore these chains from a perspective due to Athreya and Ney [2] and Nummelin [17]. This perspective, while natural in its own right, is especially appropriate from a simulation standpoint, as it essentially tries to reduce analysis of a general state space chain to that of a regenerative process.

We start, in Section 2, by defining the abstract notion of a Markov chain on general state space. Some elementary properties of such chains are then explored, and several examples of such processes are introduced.

In Section 3, we give the definition of Harris recurrence. We then follow the method of Athreya, Ney, and Nummelin in showing how a certain subclass of such processes, called strongly recurrent chains, can be embedded in a regenerative environment. Consequences of the embedding are displayed, and the examples of Section 2 are analyzed from this recurrence viewpoint. We also give a means of reducing many weakly recurrent (non strongly recurrent) chains to a strongly recurrent process. We then relate our concept of a Harris chain to the classical one (see e.g., Harris [13]) and use the relationship to prove results

concerned with guaranteeing Harris recurrence for certain types of chains.

Section 4 deals with a historical precursor of Harris recurrence known as Doeblin recurrence (Doob [9]). The relationship between Doeblin and Harris chains is investigated from the Athreya-Ney-Nummelin perspective introduced in Section 3. A result showing that a special class of Harris chains can be weakly approximated by Doeblin processes is then demonstrated.

Ergodic theory for Harris chains is dealt with in Section 5. Emphasis is placed on the relationship between ergodic behavior and certain periodicity and uniformity conditions. Some comments are also made on strong laws of large numbers for such chains that pertain to simulation of such processes.

Section 6 deals with regenerative simulation of the strongly recurrent chains defined in Section 3. The application of the regenerative method of simulation to this situation is outlined. To initiate the regenerative simulation, a certain measure is required. An optimality condition concerning the measure is considered, and a means of determining it is indicated. Also, an acceptance-rejection method intimately connected with the basic regenerative framework is given. Finally, practical simulation difficulties involved with the method are discussed.

In Section 7, it is shown that a broad class of generalized semi-Markov processes is only weakly recurrent. This motivates the development of a framework, very similar to that of Athreya, Ney, and Nummelin, for embedding such chains in a "loosely regenerative" environment. A

central limit theorem, from which confidence intervals may be constructed, is given for such processes. Finally, the necessary modifications in the methods of Section 6, for simulation of weakly recurrent chains, are outlined.

2. General State Space Markov Chains

In this section, we introduce the concept of a Markov chain (M.C.) on general state space. We then proceed to explore some elementary properties of such chains.

Let S be a non-empty set and $\mathcal{S}$ a σ-field of subsets of S. A function $P:S \times \mathcal{S} \to [0,1]$ is called a probability transition kernel if:

i) $P(x,\cdot)$ is a probability measure on $(S,\mathcal{S})$, for every x in S.

ii) $P(\cdot,B)$ is an $\mathcal{S}$-measurable function, for every $B \in \mathcal{S}$.

The n-step transition probabilities are defined through the relations

$$P^1(x,B) = P(x,B)$$

(2.1)

$$P^{n+1}(x,B) = \int_S P^n(y,B)\, P(x,dy) , \qquad n \geq 1 .$$

The preceding definitions specify the state space and finite-dimensional characteristics of the probability transition mechanism. We now turn to construction of the process.

Let $\Omega = S \times S \times \cdots$ and let $\mathcal{F}$ be the product σ-field $S \times S \times \cdots$. For $\omega = (\omega_0, \omega_1, \ldots) \in \Omega$, put $X_i(\omega) = \omega_i$. Then, for each $x \in S$, there exists a unique probability measure P_x on $(\Omega,\mathcal{F})$ such that for any $n \geq 1$ and $B_1, B_2, \ldots, B_n \in \mathcal{S}$,

$$P_x\{X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n\}$$

$$= \int_{B_1} P(x, dy_1) \int_{B_2} P(y_1, dy_2) \cdots \int_{B_n} P(y_{n-1}, dy_n) \quad .$$

Also, as a consequence of the measurability of the kernel, the map $x \to P_x(E)$ is $\mathcal{S}$-measurable, for every $E \in \mathcal{F}$ (see Revuz [20], p. 16). Hence, we can define, for an arbitrary measure μ on $(S,\mathcal{S})$, the measure P_μ on $(\Omega,\mathcal{F})$ via

$$P_\mu(E) = \int_S P_x(E)\mu(dx) \tag{2.2}$$

Clearly, for $\mu = \delta_x$ ($\delta_x(B) = 1$ or 0 depending on whether or not $x \in B$), $P_\mu = P_x$. The relation (2.2) extends to expectations in the sense that if Z is a real-valued $\mathcal{F}$-measurable function, then

$$E_\mu Z = \int_S E_x Z \, \mu(dx) \tag{2.3}$$

provided that at least one side of the equality is well-defined.

It can be readily verified that the process $\{X_n;\ n \geq 0\}$ constructed relative to $(\Omega,\mathcal{F},P_\mu)$ has the property that

(2.4) $$P_\mu\{X_{n+k} \in B|F_n\} = P^k(X_n,B)$$

P_μ a.s., where $F_n = B(X_0, \ldots, X_n)$ (the smallest σ-field generated by $X_0, \ldots, X_n$). This is just the statement of the ordinary Markov property (MP).

The process $\{X_n\}$ just constructed also has the Strong Markov property (SMP). To be precise, let $\theta_n:\Omega \to \Omega$ be the shift given by

(2.5) $$\theta_n(\omega) = (\omega_n, \omega_{n+1}, \ldots) \quad .$$

We say that $T:\Omega \to \{1, 2, \ldots\}$ is a stopping time if $\{T = n\} \in F_n$. The σ-field F_T associated with T is given by the set of events $E \in F$ such that $\{T = n\} \cap E \in F_n$ for all n. Then, the SMP states that

(2.6) $$E_\mu\{Z\circ\theta_T|F_T\} = E_{X_T} Z$$

P_μ a.s., where Z is any real-valued F-measurable function with $E_\mu|Z| < +\infty$.

The process just constructed will be called the canonical M.C. associated with P. We now introduce some examples of such processes.

(2.7) Example

Let $S = \{0, 1, 2, \ldots\}$ and suppose $P = \{P_{ij};\ i,j \in S\}$ is a stochastic matrix. Then, $\{X_n\}$ is a countable state M.C.

(2.8) Example

Consider the $\{W_n\}$ waiting time process associated with a G/G/1 queue. Let U,V be random variables with the relevant interarrival and service distributions, respectively, and suppose that $F(x) = P\{V-U \leq x\}$ has continuous density $f(x)$. The state space here is $S = [0, +\infty)$, with $\mathcal{S}$ the Borel sets on S. Also, for $B \in \mathcal{S}$

$$P(x,B) = P\{W_{n+1} \in B | W_n = x\}$$

$$= P\{(x + V_n - U_{n+1})^+ \in B\}$$

$$= \int_{B-x} f(y)dy + \delta_0(B) \int_{(-\infty,-x]} f(y)dy$$

where $B-x = \{z: z+x \in B\}$. Henceforth, by $\{W_n\}$ process, we shall mean the canonical chain associated with the above kernel.

(2.9) Example

Let $Z_1, Z_2, \ldots$ be a sequence of independent, identically distributed (i.i.d.) random variables with common distribution function F having support on $[0, 2b]$, $b > 0$, Also, suppose that F has a continuous density f. Define, for $0 < y < 2b$, a sequence $X_0, X_1, \ldots$ through

$$X_0 = y ,$$

$$X_{n+1} = |X_n - Z_{n+1}| , \qquad n \geq 0 .$$

The sequence $\{X_n\}$ can be realized as a canonical chain relative to (Ω,F,P_y), where the probability triple was constructed from $S = [0, 2b]$, S = Borel sets on S, and

$$P(x,B) = P\{|x-Z| \in B\}$$

$$= \int_{B\cap[0,x]} [f(x-y) + f(x+y)]dy + \int_{B\cap[x,2b]} f(x+y)dy \quad .$$

Feller [10], p. 208, can be consulted for further details.

(2.10) Example

A class of stochastic processes that appears to hold promise as a mathematical model of the general discrete event simulation is the generalized semi-Markov process (GSMP). A GSMP is a process on a countable state space that associates with each state a set of clock readings. When the clock with the minimal reading runs down to zero, a state transition is initiated. The process jumps to another state, where the clocks from the preceding state continue to run down, and where new clocks (including possibly the clock that initiated the jump) are set stochastically to new values. This procedure then repeats itself to generate a sample path of the GSMP.

These processes can be analyzed by studying an associated general state space Markov chain, namely the process that records the state of the GSMP together with all relevant clock readings, at the time of each transition. We refer the reader to Whitt [23] for a detailed description of the process, and for the notation that shall appear below.

The state space S of the M.C. is given by Σ, the set of all GSMP states together with all possible associated clock readings. The kernel P is

(2.11) $P((s,c), A)$

$$= p(s'; s, i^*) \prod_{i \in N_{s'}} F(a_i; s', i, s, i^*) \prod_{i \in O_{s'}} I_{[0,a_i]}(c_i^*) ,$$

where I_B is the indicator function of the set B. The pair (s,c) represents the state s of the GSMP together with its associated clock readings c. A transition to A requires entering the state $\{s'\}$, with the i'th clock associated with $\{s'\}$ set at a reading less than or equal to a_i.

The numbers $p(s'; s, i^*)$ govern state transitions of the GSMP, and represent the probability of a jump to s', given that the process was in s, and that clock i^* initiated the jump. The rest of the kernel governs clock readings. Those clocks $i \in N_{s'}$ are set stochastically according to a distribution function $F(a_i; s', i, s, i^*)$, whereas those clocks $j \in O_{s'}$, are inherited from the previous state and so must be set deterministically (hence the indicators) at the previous value c_j^*.

3. Harris Recurrent Markov Chains

We note that for Example 2.9, there exist no points in S that are visited infinitely often (i.o.), and yet one can reasonably expect that this process has "nice" ergodic behavior. Therefore, any good definition of recurrence in general state space should be free of such an assumption.

We say that a M.C. $\{X_n\}$ is Harris recurrent if there exists a tuple $(A, B, \lambda, \varphi, n)$ such that

i) $P_x\{T_A < +\infty\} = 1$, for every $x \in S$, where $T_A = \inf\{n \geq 1 : X_n \in A\}$

ii) $P^n(x,E) \geq \lambda\varphi(E)$ for all $x \in A$, measurable $E \subseteq B$, where $\lambda > 0$ and φ is a probability on B.

This notion of recurrence was first introduced and studied by Athreya and Ney [2] and Nummelin [17].

Remarks

(3.1) The second condition is a smoothness condition on the kernel. The first condition in fact guarantees that A is hit i.o. Observe that the consecutive hitting times T_n of A are given by $T_1 = T_A$, $T_n = T_{n-1} + T_1 \circ \theta_{T_{n-1}}$. Then

$$\begin{aligned} P_x\{T_2 < +\infty\} &= P_x\{T_1 < +\infty,\ T_1 \circ \theta_{T_1} < +\infty\} \\ &= E_x\{I_{\{T_1 < +\infty\}}\, P_{X_{T_1}}\{T_1 < +\infty\}\} \\ &= P_x\{T_1 < +\infty\} = 1 \end{aligned}$$

so that $T_2 < +\infty$ P_x a.s. Now, proceed inductively.

(3.2) The λ parameter is constrained by $0 < \lambda \leq 1$. This follows from

$$1 \geq P^n(x,B) \geq \lambda\varphi(B) = \lambda > 0 \quad .$$

(3.3) Note that condition ii) implies that $\varphi(\cdot)$ is absolutely continuous with respect to $P^n(x,\cdot)$ ($\varphi << P^n(x,\cdot)$) for every $x \in A$. Hence, by the Radon-Nikodym Theorem (see e.g. Royden [21]), the derivative of φ with respect to (w.r.t.) $P^n(x,\cdot)$ exists and satisfies

$$\lambda\left[\frac{d\varphi}{dP^n(x,\cdot)}\right](y) \leq 1 \tag{3.4}$$

$P^n(x,\cdot)$ a.e. in y.

(3.5) Without loss of generality, we can suppose that $A = B$ in the above definition. In other words, if there exists (A,B,λ,φ,n) satisfying the above, then there also exists a tuple $(A',A',\lambda',\varphi',n')$ satisfying i) and ii). In fact, one can take $A = A'$. See Lemma 3.1 for a proof.

(3.6) We shall call a set A, such that there exists λ, φ, n with (A,A,λ,φ,n) a Harris tuple, an A-set.

(3.7) Lemma

Suppose $\{X_n\}$ is a process with kernel P satisfying (A,B,λ,φ,n) recurrence. Then the process is also $(A,A,\lambda',\varphi',n')$ recurrent.

Proof. By condition i) and eqn. (2.4)

$$P_\varphi\{T_A < +\infty\} = \int_S P_x\{T_A < +\infty\}\, \varphi(dx) = 1$$

so that there exists an integer m such that

$$P_\varphi\{X_m \in A\} = \int_S P^m(x,A)\, \varphi(dx) > 0 \quad .$$

Then, for x in A, and measurable $E \subseteq A$, we can utilize (3.3) to get

$$\begin{aligned} P^{n+m}(x,E) &= \int_S P^m(y,E)\, P^n(x,dy) \\ &\geq \lambda \int_S P^m(y,E) \left[\frac{d\varphi}{dP^n(x,\cdot)}\right](y)\, P^n(x,dy) \\ &= \lambda'\, \varphi'(E) \end{aligned}$$

where

$$\varphi'(E) = \int_S P^m(y,E)\, \varphi(dy)/P_\varphi\{X_m \in A\}$$

$$\lambda' = \lambda P_\varphi\{X_m \in A\} \ .$$

This verifies the second condition of Harris recurrence for the tuple $(A,A,\lambda',\varphi',n+m)$. The first condition is clearly satisfied. ||

We return to our examples of Section 2.

Example 2.7 continued

Suppose $\{X_n\}$ is an irreducible recurrent M.C. Then there exists $i \in S$ such that $\{i\}$ is visited i.o. by $\{X_n\}$. Here, one puts $A = \{i\}$, $B = S$, $\lambda = 1$, $\varphi_j = P_{ij}$, $n = 1$ to get the required tuple.

Example 2.8 continued

We know that if $EV \leq EU$, $\{0\}$ is visited i.o. by the $\{W_n\}$ chain. Here, we take $A = \{0\}$, $B = S$, $\lambda = 1$, $\varphi(E) = P(0,E)$, $n = 1$.

Example 2.9 continued

Suppose that f is positive over $[0,2b]$, so that $\min\{f(x):x \in [0,2b]\} = \alpha > 0$. Then, putting $A = [0,b]$, we see that for any Borel subset E of A, and x in A,

$$P(x,E) \geq \int_E f(x+y)dy \geq \alpha m(E)$$

where m is Lebesgue measure. It's not difficult to show that A is visited i.o. and so, by putting $B = A$, $\lambda = \alpha b$, $\varphi(\cdot) = m(\cdot)/b$, $n = 1$, we see that $\{X_n\}$ is Harris recurrent.

As might be expected, the integer n of the condition ii) is closely related to the concept of periodicity for general state space chains.

We say that $\{X_n\}$ is strongly aperiodic if the chain is $(A,A,\lambda,\varphi,1)$ recurrent. To define further notions of periodicity for Harris chains, we consider

$$d_A = \gcd\{k : \exists \lambda_k, \varphi_k \text{ with } \{X_n\} (A,A,\lambda_k,\varphi_k,k) \text{ recurrent}\}$$

for A-sets A. It can be shown that d_A is in fact independent of the choice of A (Athreya and Ney [1], p. 7). It is then said that $\{X_n\}$ is periodic of period d_A if $d_A > 1$, and aperiodic if $d_A = 1$. Given that $\{X_n\}$ is of period d, the state space S can be decomposed into sets $F, C_1, \ldots, C_d$ such that

i) $F = S - \bigcup_{i=1}^{d} C_i$ and $P_x\{X_n \in F \text{ i.o.}\} = 0$ for every x in S.

ii) $C_1, \ldots, C_d$ is a cycle, meaning that

$$P(x, C_{j+1}) = 1, \qquad x \in C_j, \ 1 \leq j \leq d-1$$

$$P(x, C_1) = 1, \qquad x \in C_d$$

(see Orey [19], p. 13). We shall see in Section 5 the way in which periodicity affects the ergodic behavior of a Harris chain.

We will now indicate how the Harris condition implies nice ergodic behavior for the chain. It turns out that when a kernel P is $(A,B,\lambda,\varphi,1)$

recurrent, then one can embed the original process in one with a larger state space, such that the new process is regenerative. One then obtains "nice" ergodic behavior for the extended process, which is subsequently inherited by the original process.

Under the condition of $(A,B,\lambda,\varphi,1)$ recurrence, we can decompose P as

$$P(x,E) = \lambda\varphi(E) + (1-\lambda)\, Q(x,E) \tag{3.8}$$

for $x \in A$, E in B. The function Q is a probability transition kernel on $(S,\mathcal{S})$ and is uniquely determined by equation (3.8) for $0 < \lambda < 1$ (for $\lambda = 1$, Q is irrelevant).

Athreya, Ney, and Nummelin exploited the decomposition (3.8) in the following embedding. Let $S' = S \times \{0,1\}$, and $\mathcal{S}'$ the associated product σ-field. Define a kernel P' relative to $(S',\mathcal{S}')$ by

$$P'((x,\delta),\, E \times \{0\}) = \begin{cases} (1-\lambda)\, P(x,E), & x \notin A \\ (1-\lambda)\, Q(x,E), & x \in A \end{cases}$$

$$P'((x,\delta),\, E \times \{1\}) = \begin{cases} \lambda P(x,E), & x \notin A \\ \lambda\varphi(E \cap B), & x \in A \end{cases} \tag{3.9}$$

For a measure μ on $(S,\mathcal{S})$, define μ' on $(S',\mathcal{S}')$ by

(3.10) $$\mu'(E \times \{i\}) = \delta_0(i)\ (1-\lambda)\ \mu(E) + \delta_1(i)\lambda\ \mu(E) \quad .$$

Then, if we let $\{X_n'\} = \{(X_n, \delta_n)\}$ be the canonical M.C. corresponding to S' and $P_{\mu'}'$, it can be shown that:

i) for any μ on $(S,\mathcal{S})$, the marginal distribution (w.r.t. $P_{\mu'}'$) of the coordinate process $\{X_n\}$ of the M.C. $\{X_n'\}$ and the distribution (w.r.t. P_μ) of the original M.C. $\{X_n\}$ are identical.

ii) the random variables $\{\delta_n\}$ constitute a sequence of i.i.d. Bernoulli deviates.

Remark

(3.11) The correspondence between $P_{\mu'}'$ and P_μ means that sample path properties of $\{X_n\}$ hold under $P_{\mu'}'$ if and only if they hold under P_μ. For example, proving a strong law of large numbers for $\{X_n\}$ under $P_{\mu'}'$ means that it will also hold for $\{X_n\}$ under P_μ. Similarly,

$$P_{\mu'}'\{f(X_0, \ldots, X_n) \leq x\} \to F(x)$$

$$\text{iff} \quad P_\mu\{f(X_0, \ldots, X_n) \leq x\} \to F(x)$$

so that central limit theorems will also coincide.

Define $N_1 = \inf\{n \geq 1 : X_{n-1} \in A,\ \delta_n = 1\}$ and put $N_k = \inf\{n > N_{k-1} : X_{n-1} \in A,\ \delta_n = 1\}$ for $k > 1$. Then, for every $k \geq 1$, $N_k < +\infty$, $P_{\mu'}'$ a.s. and the sequence $\{N_k;\ k \geq 1\}$ constitute regeneration points for the

process X'_n. The regenerative property stems from the fact that

$$P'_{\mu'}\{X'_{n+N_1} \in E\} = P'_{\mu'}\{X'_{n+\theta_{N_1}} \in E\}$$

$$= P'_{X'_{N_1}}\{X'_n \in E\}$$

$$= P'_{\hat{\varphi}}\{X'_n \in E\}$$

where $\hat{\varphi}(E \times \{i\}) = \delta_1(i)\,\varphi(E)$. The fact that X'_n is regenerative means that one can expect strong laws of large numbers and central limit theorems to be valid, under reasonably mild conditions. For example, the classical regenerative theory (see e.g. Smith [22]) shows that if $E'_{\hat{\varphi}} N_1 < +\infty$, then the strong law

$$\frac{1}{n+1}\sum_{k=0}^{n} f(X_k) \to E'_{\hat{\varphi}}\Big(\sum_{j=0}^{N_1-1} f(X_j)\Big)\Big/E'_{\hat{\varphi}} N_1 \tag{3.12}$$

holds $P'_{\mu'}$ a.s., provided that the numerator of the ratio on the right hand side is finite. This will always be the case if $f = I_A$, and then (3.12) becomes

$$\frac{1}{n+1}\sum_{k=0}^{n} I_A(X_k) \to E'_{\hat{\varphi}}\Big(\sum_{j=0}^{N_1-1} I_A(X_j)\Big)\Big/E'_{\hat{\varphi}} N_1 \quad . \tag{3.13}$$

Denote the limit by $\pi(A)$. It can be shown (Athreya and Ney [2], p. 499) that the set function $\pi(A)$ is a probability measure satisfying

(3.14) $$\pi(E) = \int_S P(x,E)\, \pi(dx) \quad .$$

A measure satisfying (3.14) is called an invariant measure for P. Relation (3.13) is then a classical ergodic theorem stating that the time average of the process tends to the spatial average.

Remarks

(3.15) Using a slightly more sophisticated argument, it can be shown (Athreya and Ney [2], p. 499) that every Harris chain possesses a σ-finite invariant measure. This invariant solution is unique among all σ-finite measures, up to a positive multiplicative constant.

(3.16) Relation (3.13) shows that

$$\varphi(E) = E'_{\hat{\varphi}}(I_E(X_0)) \leq E'_{\hat{\varphi}}\Big(\sum_{j=0}^{N_1-1} I_E(X_j)\Big) = \pi(E)\, E'_{\hat{\varphi}}\ N_1$$

and hence $\varphi \ll \pi$. This will also be true of general Harris chains.

Let us say that a chain is strongly recurrent if it is $(A,B,\lambda,\varphi,1)$ recurrent. The following example shows that not all Harris chains are strongly recurrent.

(3.17) Example

Let $\{X_n\}$ be an $(A,A,\lambda,\varphi,1)$ M.C. on (S,S) with kernel P. Then, the M.C. $\hat{X}_n = (X_n, X_{n+1})$ has state space $S \times S$ and kernel

$$\hat{P}((x,y),\ E_1 \times E_2) = \delta_y(E_1)\ P(y,\ E_2) \quad .$$

We claim that $\hat{P}$ is Harris. Observe that if $E_1,\ E_2 \subseteq A$, then for $(x,y) \in A \times A$,

$$\hat{P}^2((x,y),\ E_1 \times E_2) = \int_{E_1} P(y,\ dz)\ P(z,\ E_2) \geq \lambda^2\ \varphi(E_1)\ \varphi(E_2) \quad .$$

Furthermore, it is clear that $A \times A$ is visited i.o. by the process $\hat{X}$, and so $\hat{X}$ is $(A \times A,\ A \times A,\ \lambda^2,\ \varphi \times \varphi,\ 2)$ recurrent.

Now, suppose that $\hat{P}$ is a strongly recurrent kernel, say $(B_1,\ B_2,\ \alpha,\ \psi,\ 1)$ recurrent. Then, since $\psi << P((x,y),\ \cdot)$ (Remark (3.3)) for some (x,y), ψ must be supported by $\{y\} \times S$. Under the assumption of strong recurrence the chain must be able to pass from B_1 into $\{y\} \times S$ in one step. Consequently, $B_1 \subseteq S \times \{y\}$. However, it is clear that in general $\hat{X}$ will not hit a B_1 of this form i.o. Hence, $\hat{X}$ will not, in general, be strongly recurrent.

Although Harris chains are not always strongly recurrent, one can often modify the process, in a natural way, so that the resulting chain is strongly recurrent.

Given a kernel P, we say that $\hat{P}$ is the randomized version of P if it has the form $(0 < p < 1)$

$$\hat{P}(x,E) = \sum_{k=1}^{\infty} p(1-p)^{k-1}\ P^k(x,E) \quad .$$

(3.18) Lemma

Let P be a Harris kernel with invariant probability π. Then, $\hat{P}$ is a strongly aperiodic Harris kernel.

Proof: Suppose that P is $(A, A, \lambda, \varphi, n)$ recurrent. Then, for $x \in A$ and $E \subseteq A$,

$$\hat{P}(x,E) \geq p(1-p)^{n-1} P^n(x,E)$$

$$\geq p(1-p)^{n-1} \lambda\varphi(E)$$

so that $\hat{P}$ satisfies the second condition of $(A, A, p(1-p)^{n-1}\lambda, \varphi, 1)$ recurrence. We also need to show that

$$\hat{P}_x\{T_A < +\infty\} = 1$$

for every $x \in S$. With the existence of π, this fact is proved in the Appendix. ||

Remark

(3.19) The $\hat{P}$ chain is obtained from the P chain by sampling from the original P chain according to a sequence of i.i.d. Bernoulli random variables of parameter p. This motivates use of the term 'randomized' version.

We shall now describe a notion of recurrence that is virtually equivalent to that described above. This definition requires, however, a separability assumption on the σ-field $\mathcal{S}$. Therefore, we will assume throughout the remainder of this paper that $\mathcal{S}$ is countably generated.

Remarks

(3.20) The separability assumption is automatically satisfied if $\mathcal{S}$ is the Borel σ-field associated with a separable metric space S. This is because the σ-field is generated by the open balls with rational radii centered at points lying in some countable dense subset of S.

(3.21) The separability assumption can be avoided to some extent by using the concept of admissible σ-fields (see Orey [19], Doob [9]).

A chain $\{X_n\}$ is said to be Harris ν-recurrent if there exists a σ-finite measure ν on $(S,\mathcal{S})$ such that $\nu(E) > 0$ implies

$$P_x\{T_E < +\infty\} = 1$$

for every x in S (see e.g., Harris [13]).

We claim that under the assumption of separability of $\mathcal{S}$, Harris ν-recurrence and Harris (A,B,λ,φ,n) recurrence apply to exactly the same class of chains. First, suppose that the chain $\{X_n\}$ is (A,B,λ,φ,n) recurrent. Then, the process is clearly Harris φ-recurrent.

For the converse, we use the concept of C-sets (Orey [19], p. 7). By way of introduction, we recall that for every positive integer n, we can decompose $P^n(x,\cdot)$ into its ν-singular and ν-absolutely continuous parts, to yield

$$P^n(x,E) = \int_E p^n(x,y)\ \nu(dy) + P_S^n(x,E) \tag{3.22}$$

where $P_S^n(x,\cdot)$ is the ν-singular part of the decomposition. A set $C \in S$ is said to be a C-set if $\nu(C) > 0$ and if there exists m such that

$$\inf_{(x,y) \in C \times C} p^n(x,y) = \lambda > 0 \quad . \tag{3.23}$$

Let $0 < \nu(E) < +\infty$. Then the separability of S allows one to find a well enough behaved version of $p^n(x,y)$ that a C-set C contained in E can be produced. Observe that $\nu(C) > 0$ so that $\{X_n\}$ visits C i.o. Furthermore, for x in C and B in C,

$$P^n(x,B) \geq \int_E p^n(x,y)\ \nu(dy) \geq \lambda\nu(B)$$

so that $\{X_n\}$ is $(C, C, \lambda\nu(C), \nu(\cdot)/\nu(C), n)$ recurrent.

Remarks

(3.24) The concept of Harris ν-recurrence is in fact the classical definition of Harris recurrence as given, for example, in Orey [19] and Revuz [20].

(3.25) Note that $\{X_n\}$ is always φ-recurrent under the φ of an $(A, B, \lambda, \varphi, n)$ tuple. However, given that $\{X_n\}$ is ν-recurrent, there does not generally exist A, B, λ, n such that $\{X_n\}$ is (A, B, λ, ν, n) recurrent.

We now give some indications of the generality of Harris recurrence. First, we define the notation, for μ a measure on $(S,\mathcal{S})$,

$$\mu f = \int_S f(y)\, \mu(dy) \quad .$$

(3.26) Lemma

Let $\{X_n\}$ be a M.C. with the following ergodic behavior. Suppose that there exists a probability π such that if f is a bounded $\mathcal{S}$-measurable function, then

$$(3.27) \qquad \frac{1}{n+1} \sum_{k=0}^{n} f(X_k) \to \pi f \ , \qquad P_x \text{ a.s.}$$

for every x in S. Then $\{X_n\}$ is Harris recurrent.

Proof. For $E \in \mathcal{S}$, take $f = I_E$. Then,

$$\frac{1}{n+1} \sum_{k=0}^{n} I_E(X_k) \to \pi(E) \ , \qquad P_x \text{ a.s.}$$

so that if $\pi(E) > 0$, then $P_x\{\sum_{k=0}^{\infty} I_E(X_k) = +\infty\} = 1$ for every x. But

$$\{\sum_{k=0}^{\infty} I_E(X_k) = +\infty\} = \{X_k \in E \text{ i.o.}\}$$

so $\{X_n\}$ is Harris π-recurrent.

Remark

(3.28) Above, we showed that (3.27) is sufficient for Harris rucurrence. In Section 5, we will shows it is also necessary, in some sense.

We will now apply Lemma 3.26 to obtain Harris recurrence for a fairly general class of chains. Let S be a separable metric space, with $\mathcal{S}$ its Borel sets. Put

$b\mathcal{S}$ = class of bounded $\mathcal{S}$-measurable functions

bC = class of bounded functions continuous with respect to metric topology on S.

One can regard P as an operator on $b\mathcal{S}$ through the relation

$$(Pf)(x) = \int_S P(x,dy)\, f(y) .$$

We will say that P is a Feller kernel if P maps functions from bC back into bC.

(3.29) Lemma

P is Feller if and only if $x_n \to x$ implies $P(x_n,\cdot) \Rightarrow P(x,\cdot)$ ($\Rightarrow$ is weak convergence of probabilities).

Proof. Now, $P(x_n,\cdot) \Rightarrow P(x,\cdot)$ if and only if for every $f \in bC$

$$\int_S f(y)\, P(x_n,dy) \to \int_S f(y)\, P(x,dy) \qquad \text{i.e., } (Pf)(x_n) \to (Pf)(x) .$$

But this is precisely the notion of continuity for Pf. The boundedness of Pf is trivial.

Remark.

(3.30) It is easily checked that the kernels P of Examples 2.7 - 2.10 are Feller.

(3.31) Lemma

Let S be a compact metric space, and suppose P is Feller on $(S,\mathcal{S})$. Also, suppose that P possesses a unique invariant probability π that assigns zero mass to ∂A, for any $A \in \mathcal{S}$ (∂A is the boundary of A). Then, P is a Harris kernel.

Remark

(3.32) Every compact metric space is separable (see e.g., Copson [6], p. 77).

Proof. We first define the random measures

$$M_n(E,\omega) = \frac{1}{n+1} \sum_{k=0}^{n} I_E(X_k(\omega))$$

We next employ a theorem of Breiman [4], which states that under our hypotheses

$$\frac{1}{n+1} \sum_{k=0}^{n} f(X_k(\omega)) = \int_S f(y)\, M_n(dy, \omega) \to \pi f$$

P_x a.s., for every x in S and $f \in bC$. Now bC is a separable metric space (Remark 3.33) under the sup norm metric so that there exists a countable family C of functions dense in bC. Then $P_x(\Gamma) = 1$ where

$$\Gamma = \{\omega: \int_S f(y)\, M_n(dy, \omega) \to \pi f, \ \forall\ f \in C\} \quad .$$

For $\omega \in \Gamma$, $M_n(\cdot,\omega) \Rightarrow \pi(\cdot)$ (Remark 3.34) and so, because $\pi(\partial A) = 0$ for all Borel sets A, we get (Billingsley [3], p. 12)

$$M_n(E,\omega) \to \pi(E)$$

for every E in S. Consequently,

$$\frac{1}{n+1} \sum_{k=0}^{n} I_E(X_k) \to \pi(E) \qquad P_x \text{ a.s.}$$

for every x in S, which implies that $\{X_n\}$ is Harris π-recurrent (see Lemma 3.26).

Remarks

(3.33) Let $\{s_n\}$ be dense in S and let $\{r_n\}$ be an enumeration of the positive rationals. Then, putting

$$g_{m,n}(x) = [r_m - \rho(x,s_n)]^+$$

the family $F = \{g_{m,n}\}$ is a countable family of functions separating points in S. Then, every function in bC can be uniformly approximated by a polynomial, with rational coefficients, in a finite number of functions of F (Royden [21], p. 175). This latter family is countable.

(3.34) Let Q_n, Q be probabilities and suppose $Q_n f \to Qf$ for all f in some subset C which is dense under the sup norm $\|\cdot\|$ in bC. Then $Q_n \Rightarrow Q$ since for $f \in bC$, there exists f_n in C such that $\|f_n - f\| \to 0$ and

$$|Q_m f - Qf| \leq |Q_m f - Q_m f_n| + |Q_m f_n - Qf_n| + |Qf_n - Qf|$$

$$\leq 2\|f_n - f\| + |Q_m f_n - Qf_n|$$

which tends to zero by first letting $m \to \infty$, and then letting $n \to +\infty$.

(3.35) Note that every Feller kernel on compact metric space possesses an invariant probability (Karr [14], p. 45) so that the hypotheses of Lemma 3.31 really entail only unicity.

(3.36) An easy sufficient condition for $\pi(\partial A) = 0$ is that there exists n such that $P^n(x,\cdot) << \mu$ for every x, where $\mu(\partial A) = 0$.

Example 2.9 continued

As an easy application of Lemma 3.31, note that P is clearly Feller. In addition, $P(x,\cdot) << m$ for every x in S, and so $\pi(\partial A) = 0$ (Remark 3.36). Unicity of π follows as a consequence of $m(E) > 0$ implying $P^2(x,E) > 0$.

4. Doeblin Recurrent Markov Chains

Historically, the study of Doeblin chains preceded that of Harris processes. We say that a chain $\{X_n\}$ is Doeblin recurrent if there exists a tuple (τ, γ, η, n) consisting of a probability τ, reals $\gamma < 1$ and $\eta > 0$, and an integer n such that, for $E \in S$,

$$\tau(E) > \gamma \quad \text{implies} \quad P^n(x,E) > \eta \tag{4.1}$$

for all $x \in S$. Fossett ([11], p. 11) has shown that this condition is equivalent to that of Doob. Doob's condition states that $\{X_n\}$ is a chain with a tuple (ψ, ε, k) consisting of a finite measure ψ, a positive ε, and an integer k such that for E

$$\psi(E) < \varepsilon \quad \text{implies} \quad P^k(x,E) < 1-\varepsilon \tag{4.2}$$

for all $x \in S$. The measures ψ and τ in the two definitions are equivalent in the sense that each can play the role of the other (provided ψ is normalized to a probability).

The concept of Doeblin recurrence is closely related to that of Harris recurrence. Neveu ([16], p. 185) noted that if a process is $(S, \mathcal{S}, \lambda, \varphi, n)$ recurrent, then it is Doeblin recurrent. For if we choose ε so that $0 < \varepsilon < 1$, then $\varphi(E) > \varepsilon$ implies

$$P^n(x,E) \geq \lambda\varphi(E) > \lambda\varepsilon \tag{4.3}$$

for every x in S, so that $(\varphi, \varepsilon, \lambda\varepsilon, n)$ is a legitimate Doeblin tuple.

However, not every Doeblin process is Harris recurrent.

(4.4) Example

Take $S = \{1, \ldots, n\}$ and let P be a stochastic matrix on S. Then, if we put $\tau_i = 1/n$ for all i, we see that $\tau(E) > (n-1)/n$ implies that $E = S$, and so $P(i,E) = 1$. So every finite-dimensional stochastic matrix is Doeblin.

On the other hand, let ν be any σ-finite measure on S. Then there exists j such that $\nu_j > 0$, and so, for $\{X_n\}$ to be a ν-recurrent chain, we must have

$$P_i\{T_j < +\infty\} = 1$$

for all $i \in S$. So $\{X_n\}$ must be an irreducible chain in order that P be Harris recurrent (this is also sufficient, by Example 2.7, see Section 3).

Example 4.4 points to the difficulty, namely that we must restrict attention to irreducible chains in order to establish a proper connection between Harris and Doeblin chains. We first discuss the appropriate generalization of irreducibility to a general state space S.

A set $F \in S$ is called a transient set if

$$\lim_{n \to \infty} P^n(x,F) = 0$$

for every x in S. Doob showed ([9], p. 210) that for a Doeblin chain, there exists a decomposition of S into disjoint ergodic sets $E_1, E_2, \ldots$ and a transient set $F = S - \cup E_i$. An ergodic set is a set E for which there exists a probability π_E on E with

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} P^k(x,B) = \pi_E(B) \tag{4.5}$$

for every x in E and B in E. The ergodic sets determined by the decomposition are unique, provided points in transient sets are neglected.

The way to generalize irreducibility to this setting is to suppose that there exists a single ergodic set. This is in fact equivalent to assuming that $\{X_n\}$ has a unique invariant measure (Doob [9], p. 214).

(4.6) Lemma

Let $\{X_n\}$ be a Doeblin chain with single ergodic set. Then, $\{X_n\}$ is Harris recurrent.

Proof. Let π be the unique invariant probability of $\{X_n\}$. Then we can use Theorem 6.2 of Doob [9] (p. 220) to conclude that if $\pi|f| < +\infty$, then

$$\frac{1}{n+1} \sum_{k=0}^{n} f(X_k) \to \pi f$$

P_x a.s., for every x in S. Application of Lemma 3.26 then yields the result. ||

Given that a Doeblin chain is Harris, how does the process fit into the Harris framework? The following two lemmas address this question.

(4.7) Lemma

A chain $\{X_n\}$ is Doeblin recurrent with a single ergodic set if and only if the chain is $(A, A, \lambda, \varphi, n)$ recurrent, and there exists an integer m such that

$$\sup_{x \in S} P_x\{T_A > m\} < 1 \quad .$$

(4.8) Lemma

A process X_n is an aperiodic Doeblin chain with single ergodic set if and only if the chain $\{X_n\}$ is $(S, S, \lambda, \varphi, n)$ recurrent.

We defer the proofs of these results to Section 5.

Remark

(4.9) Chains of the type described in Lemma 4.7 correspond precisely to the uniformly ν-recurrent chains of Orey [19]. This fact can be proved by employing C-sets.

Suppose that a chain is $(S,\mathcal{S},\lambda,\varphi,n)$ recurrent. Then the tuple $(\varphi,\varepsilon,\lambda\varepsilon,n)$ is a legitimate Doeblin tuple (see (4.3)) so that φ can be used in the role of τ (see (4.1)). The converse, however, is false, as the following example shows.

(4.10) Example

Let $S = \{0, 1, 2, \ldots\}$ and let P be given by

$$P(i,E) = \frac{1}{2}\,\delta_0(E) + \frac{1}{2}\,\delta_{i+1}(E) \quad .$$

This chain is clearly $(S,\mathcal{S},1/2,\delta_0,1)$ recurent, and it is easily checked that $\pi_i = 1/2^{i+1}$ is an invarant probability for P. Furthermore, $\pi(E) > 1/2$ implies that $0 \in E$, so that

$$P(i,E) \geq 1/2$$

for all i. Hence, $(\pi, 1/2, 1/2, 1)$ is a legitimate Doeblin tuple.

On the other hand, there can exist no λ and n such that P is $(S, \mathcal{S}, \lambda, \pi, n)$ recurrent. For if such λ and n existed, then we would have

$$P^n(i,j) \geq \lambda\pi_j > 0$$

for every i and j. But this contradicts the fact that $P^n(i, i+n+1) = 0$.

This shows that although π is a legitimate τ measure, it is not a legitimate φ measure.

We close this section with a weak convergence theorem for Doeblin chains. To be precise, let S be a complete separable metric space with $\mathcal{S}$ its Borel σ-field. Then, by defining the product topology on $\Omega = S \times S \times \cdots$, Ω itself becomes a complete separable metric space. In addition, it can be shown (Dellacherie and Meyer [8], p. 11) that the σ-field $\mathcal{F}$ is precisely the Borel σ-field under this topology.

We shall say that a kernel P is strong Feller if $P(\cdot,E)$ is a continuous function for each $E \in \mathcal{S}$. Our result is that if P_μ is a probability on $(\Omega,\mathcal{F})$ constructed from a Harris recurrent strong Feller kernel P, then one can find a sequence of probabilisties P_n such that

i) P_n is of the form $P^{(n)}_{\mu_n}$, where $\mu_n \Rightarrow \mu$ and $P^{(n)}$ are Doeblin kernels.

ii) $P_n \Rightarrow P_\mu$ on $(\Omega,\mathcal{F})$, i.e., for every bounded function h continuous w.r.t. the topology on Ω,

$$\int_\Omega h(\omega)\, P_n(d\omega) \to \int_\Omega h(\omega)\, P(d\omega) \quad .$$

We start by recalling that since P is ν-recurrent, $\nu(A) > 0$ implies that A is hit i.o. P_x a.s. For such an A, let us denote the consecutive hitting times by $T_1, T_2, \ldots$. Then, the sequence $\{X_{T_i}\}$ is a M.C. with kernel

$$(4.11)\quad P_A(x,E) = P(x, E \cap A) + \int_{A^c} P(x, dy)\, P(y, E \cap A)$$

$$+ \int_{A^c}\int_{A^c} P(x, dy)\, P(y, dz)\, P(z, E \cap A) + \cdots$$

This chain is called the "process on A". Orey ([18], [19]) showed that for a ν-recurrent P, there exists a sequence A_n of sets increasing to S (i.e., $A_n \subseteq A_{n+1}$, $\cup A_n = S$) such that P_{A_n} is Doeblin on A_n (in fact, uniformly ν-recurrent). We construct P_n by noting first that since $A_n \nearrow S$, $\mu(A_n) \to 1$ so that we may suppose $\mu(A_n) > 0$ for each n. Then put

$$P^{(n)}(x,E) = P_{A_n}(x,E)$$

$$\mu_n(E) = \mu(E \cap A_n)/\mu(A_n)$$

and let $P_n = P^{(n)}_{\mu_n}$ be the corresponding measure on (Ω,F). We now employ a weak convergence theorem of Karr [14] that states that to show $P_n \Rightarrow P_\mu$, it is sufficient to prove that

i) $\mu_n \Rightarrow \mu$

ii) $x_n \to x$ in S implies $P^{(n)}(x_n,\cdot) \Rightarrow P(x,\cdot)$.

Clearly, i) is trivial. As for ii), note that for $f \in bC$,

$$|P^{(n)} f(x_n) - Pf(x)|$$

$$\leq \left|P^{(n)} f(x_n) - \int_{A_n} f(y)\, P(x_n, dy)\right|$$

$$+ \left|\int_{A_n} f(y)\, P(x_n, dy) - Pf(x_n)\right| + |Pf(x_n) - Pf(x)| .$$

It is easily checked that the first term on the right is bounded by $\|f\|\, P_{x_n}\{T_{A_n} > 1\} = \|f\|\, P(x_n, A_n^c)$ (see (4.11)). The second term is bounded by the same quantity. So, fixing m for the moement, we see that for $n \geq m$

$$|P^{(n)} f(x_n) - Pf(x)| \leq 2\|f\|\, P(x_n, A_m^c) + |Pf(x_n) - Pf(x)|$$

$$\to 2\|f\|\, P(x, A_m^c)$$

as $n \to \infty$, since P is strong Feller. Then, let $m \to +\infty$ to complete the proof of ii).

Remark

(4.12) Karr [14] also gives conditions under which the invariant measures of a sequence of M.C.'s tends to an invariant measure for the limit.

5. Ergodic Behavior of Harris Chains

The ergodic behavior of Harris chains is basically determined by periodicity considerations and by certain types of uniformity conditions on the $(A, B, \lambda, \varphi, n)$ tuple. We recall that every Harris chain has a unique invariant measure (Remark 3.15). Assuming now that this measure is a probability π, our first goal is to examine convergence to π of the iterates $P^1, P^2, \ldots$ of the transition kernel.

Then, if $\{X_n\}$ is an aperiodic Harris chain

$$(P^n f)(x) \to \pi f \tag{5.1}$$

for every x in S and f in bS. In fact, a stronger result holds. Let $\|\cdot\|$ be the L^∞ norm on bS. Then, there exist constants $c_n \downarrow 0$ (independent of f) such that

$$|(P^n f)(x) - \pi f| \leq c_n \|f\| \quad . \tag{5.2}$$

If the process is periodic, then one gets convergence of the averages, i.e., there exist constants $c_n \downarrow 0$ such that

$$\left| \frac{1}{n+1} \sum_{k=0}^{n} (P^k f)(x) - \pi f \right| \leq c_n \|f\| \quad . \tag{5.3}$$

When no invariant probability exists, one can still prove that if P is an aperiodic Harris kernel, then there exist constants $c_n \downarrow 0$ such that

$$|\mu_1 P^n f - \mu_2 P^n f| \leq c_n \|f\|$$

where μ_1, μ_2 are any two initial probabilities on (S,S). Here the operator μP^n is given by

$$\mu P^n f = \int_S \mu(dx)\ (P^n f)(x) \ .$$

Proofs for these results can be found in Athreya and Ney [1] or Orey [19].

We can get even stronger results than (5.2) by imposing uniformity conditions. Athreya and Ney ([2], p. 496) consider a strongly aperiodic $(A, A, \lambda, \varphi, 1)$ process with the additional restriction that there exist an integer m such that

$$\sup_{x \in S} P_x\{T_A > m\} < 1 \ .$$

They show that for such a chain, there exists an invariant probability π and constants b and ρ (with $0 \leq \rho < 1$) such that

$$\|P^n f - \pi f\| \leq b\rho^n \|f\| \ . \tag{5.4}$$

Remark

(5.5) We recall from (4.5) that all Doeblin chains with single ergodic set have a unique invariant probability. If such a chain is also aperiodic, then Doob [9] has shown that there exist constants and ρ (with $0 \leq \rho < 1$) such that

$$|P^n(x,E) - \pi(E)| \leq \gamma\rho^n \tag{5.6}$$

for every x in S, $E \in \mathcal{S}$.

The similarity between (5.4) and (5.5) provides the key to the proofs of Lemmas 4.7 and 4.8. The randomized kernel $\hat{P}$ will also be used here. Note that if π is invariant for P, then it is invariant for $\hat{P}$.

On the other hand, suppose that π is invariant for $\hat{P}$. First of all, a glance at the proof in the Appendix shows that the assumption of an invariant probability for P made in Lemma 3.18 was asserted only to guarantee existence of an invariant probability for $\hat{P}$. Hence, given that P is Harris, $\hat{P}$ is also, and so $\hat{P}$ has a unique invariant probability, namely π. But $\pi\hat{P} = \pi$ implies that $(\pi P)\hat{P} = \pi P$, and so unicity of invariant probability for $\hat{P}$ implies that $\pi P = \pi$. Hence, provided P is Harris, invariant probabilities for P and $\hat{P}$ coincide.

Proof of Lemma 4.7

$\Rightarrow$) Since $\{X_n\}$ is Doeblin with single ergodic set, P is Harris with invariant probability π, and hence $\hat{P}$ is Harris (Lemma 3.18). Now, supposing that P offers a (τ, γ, η, m) Doeblin tuple, we have $\tau(E) > \gamma$ implies $P^m(x,E) > \eta$ so that

$$\hat{P}(x,E) \geq p(1-p)^{m-1} P^m(x,E) \geq p(1-p)^{m-1} \eta$$

and hence $\hat{P}$ is Doeblin with $(\tau, \gamma, p(1-p)^{m-1}\eta, 1)$ tuple. Now, by Lemma 3.18, $\hat{P}$ is an aperiodic Harris chain. Also, because $\hat{P}$ is Doeblin with unique invariant probability, it must be that it has a single ergodic set. Summarizing, $\hat{P}$ is an aperiodic Doeblin kernel with single ergodic set.

So, we can apply Remark 5.5 to conclude existence of a γ and ρ such that

$$|\hat{P}^n(x,E) - \pi(E)| \leq \gamma\rho^n \quad .$$

Now, P is Harris so there exists an $(A, A, \lambda, \varphi, k)$ tuple with $\pi(A) > 0$. Then, we can find n such that

$$\hat{P}^n(x,A) \geq \pi(A)/2$$

i.e.,

$$\sum_{j=n}^{\infty} a_j P^j(x,A) \geq \pi(A)/2 \tag{5.7}$$

where (a_j) is a probability. Choose ℓ so that

$$\sum_{j=\ell+1}^{\infty} a_j < \pi(A)/4 \quad .$$

Then, (5.7) becomes

$$\sum_{j=n}^{\ell} a_j P^j(x,A) \geq \pi(A)/4 \quad .$$

Letting $a = \max\{a_j;\ n \le j \le \ell\}$, we get the inequality

$$\sum_{j=n}^{\ell} P^j(x,A) \ge \pi(A)/4a \quad . \tag{5.8}$$

Putting $S_j = \{x : P^j(x,A) \ge \pi(A)/4a\ell\}$, we see that $\bigcup_{j=1}^{\ell} S_j = S$, which clearly implies that

$$P_x\{T_A > \ell\} \le 1 - \pi(A)/4a\ell \quad .$$

= The uniformity condition on P requires that

$$\inf_x P_x\{T_A \le m\} = \delta > 0$$

and hence

$$\sum_{k=1}^{m} P^k(x,A) \ge \delta$$

for every x in S. If we now form the sets $S_j = \{x : P^j(x,A) \ge \delta/m\}$, we observe that $\bigcup_{j=1}^{m} S_j = S$.

Then, for $x \in S_j$, we obtain the lower bound

$$P^{m+n}(x,E) \geq \int_A P^{m+n-j}(y,E)\ P^j(x,dy)$$

$$\geq \int_A P^j(x,dy) \int_A P^n(y,dz)\ P^{m-j}(z,E)$$

$$\geq \int_A P^j(x,dy) \int_A \lambda\varphi(dz)\ P^{m-j}(z,E)$$

$$\geq \frac{\lambda\delta}{m} \int_A \varphi(dz)\ P^{m-j}(z,E) \quad ,$$

the assumption of $(A, A, \lambda, \varphi, n)$ recurrence having been invoked in the third inequality in the chain.

We now define the probability

$$\tau(E) = \frac{1}{m} \sum_{k=1}^{m} \int_A \varphi(dz)\ P^{m-j}(z,E) \quad .$$

The inequality $\tau(E) \geq 1-\gamma$, where $0 < \gamma < \min(1/m, 1/2)$, then yields

$$\frac{1}{m} \int_A \varphi(dz)\ P^{m-j}(z,E) \geq \frac{1}{m} - \gamma$$

for $j = 1, \ldots, m$. Consequently, it must be that

$$P^{m+n}(z,E) \geq \frac{\lambda\delta}{m}(1-\gamma m) = \delta^* \ .$$

We may therefore conclude that $(\tau, 1-\lambda, \delta^*, m+n)$ is a legitimate Doeblin tuple.

Proof of Lemma 4.8

$\Rightarrow$) By Lemma 4.6, $\{X_n\}$ is Harris recurrent so there exists an $(A, A, \lambda, \varphi, n)$ tuple with $\pi(A) > 0$. The aperiodicity of the chain allows one to use (5.5) to conclude that there exists an integer m such that

$$P^m(x,A) > \pi(A)/2 = \delta .$$

Then, for every x in S and E in A, we get

$$P^{n+m}(x,E) = \int_S P^n(y,E)\ P^m(x,dy)$$

$$\geq \int_A P^n(y,E)\ P^m(x,dy)$$

$$\geq \lambda\varphi(E)\ P^m(x,A) \geq \lambda\delta\varphi(E)$$

so that $\{X_n\}$ is $(S, S, \lambda\delta, \varphi, n+m)$ recurrent.

$\Leftarrow$) For such a chain, the uniformity condition of Lemma 4.7 is trivially satisfied, and hence the process must be Doeblin recurrent with single ergodic set. Also, for any $k \geq 0$,

$$P^{n+k}(x,E) = \int_S P^k(y,E)\ P^n(x,dy)$$

$$\geq \lambda \int_S P^k(y,E)\ \varphi(dy)$$

so that $\{X_n\}$ is $(S, \mathcal{S}, \lambda, \varphi P^k, n+k)$ recurrent. This clearly implies that $\{X_n\}$ is aperiodic.

Let $\{X_n\}$ be an aperiodic Harris chain. We have seen that the uniformity condition manifests itself in the global convergence inherent in (5.4). We now examine the differing nature of the mixing conditions.

Put $\mathcal{F}^n = \mathcal{B}(X_n, X_{n+1}, \ldots)$. Then, if $\{X_n\}$ is an aperiodic Harris chain with invariant probability π, we have (Orey [19])

$$\lim_{k \to \infty} \sup_{B \in \mathcal{F}^{m+k}} |P_\pi(A \cap B) - P_\pi(A)\, P_\pi(B)| = 0$$

for any $A \in \mathcal{F}_m$. On the other hand, if the chain also satisfies the uniformity condition, then by Lemmas 4.7 and 4.8, $\{X_n\}$ is an aperiodic Doeblin chain with single ergodic set. The process then automatically possesses an invariant probability π, and there exists constants γ and ρ $(0 \leq \rho < 1)$ such that

$$\sup_{A \in \mathcal{F}_m, B \in \mathcal{F}^{m+k}} \frac{|P_\pi(A \cap B) - P_\pi(A)\, P_\pi(B)|}{P_\pi(A)} \leq \gamma \rho^k$$

(Billingsley [3], p. 168), where we interpret $0/0 = 0$.

We now turn to the question of strong laws for Harris chains. Let $\{X_n\}$ be a Harris process possessing invariant probability π. Then, it can be shown that if $\pi|f| < +\infty$,

$$\frac{1}{n+1}\sum_{k=0}^{n} f(X_k) \to \pi f \tag{5.9}$$

P_x a.s. for every x in S (see Glynn [12]). This has the important implication for simulation that one can estimate πf by forming sample averages along a single sample path of the process.

This is in contrast to the situation that occurs when one forms m independent replicates of the process, and simulates to time n along each replicate. This gives a data sample

$$\begin{array}{ccc} X_1^1, & X_2^1, & \ldots, X_n^1 \\ X_2^1, & X_2^2, & \ldots, X_n^2 \\ \vdots & \vdots & \vdots \\ X_m^1, & X_m^2, & \ldots, X_n^m \end{array}$$

Here, the estimate is formed by taking sample averages along each replicate, and averaging the replicates. The hope is that

$$\lim_{n \to \infty} \lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{m} \left(\frac{1}{n} \sum_{j=1}^{n} f(X_j^k) \right)$$

converges to πf, P_x a.s. for every x in S. But this estimate tends to

$$\lim_{n \to \infty} E_x \left(\sum_{k=1}^{n} \frac{f(X_k)}{n} \right) \tag{5.10}$$

and this is <u>not</u> in general equal to πf for a Harris chain.

(5.11) Example

Let P be the transition kernel of the $\{W_n\}$ chain of Example 2.2, and suppose that we modify it to

$$\hat{P}(n,E) = \frac{1}{2}\delta_0(E) + \frac{1}{2}\delta_{n+1}(E) , \qquad n \geq 1$$

$$\hat{P}(x,E) = P(x,E) , \qquad x \notin \{1, 2, \ldots\} .$$

Clearly, the $\hat{P}$ chain is still Harris, since $\{0\}$ is still hit i.o. Also, the modification does not disturb the invariant probability π. We now introduce a non-negative function f such that $\pi f \leq 1$ and modify it to

$$\hat{f}(n) = 2^{n+1} , \qquad n \geq 1$$

$$\hat{f}(x) = f(x) , \qquad x \notin \{1, 2, \ldots\} .$$

Then, clearly $\hat{E}_1 \hat{f}(X_n) \geq \hat{f}(n+1) \hat{P}_1\{X_n = n+1\} = 2$ so that (5.10) must be greater than $\pi\hat{f}$.

The reason that the convergence problem appeared in Example 5.11 lies in the fact that $\hat{f}$ was not bounded (compare with (5.3)). However, by using properties of f-regular measures, results of Nummelin [17] imply that

$$\frac{1}{n}\sum_{k=1}^{n} E_x f(X_k) \to \pi f$$

for π a.e. x in S. Of course, the simulator, before commencing the simulation, does not know π, and hence the above result is only mildly reassuring to those that wish to simulate across replicates.

6. Strongly Recurrent Harris Chains

In this section, we continue our study of $(A, B, \lambda, \varphi, 1)$ chains. In particular, we shall concentrate on simulation questions related to such processes.

Let us start by noting that such processes include all chains which hit a given point x_o i.o. for we can take $A = \{x_0\}$, $B = S$, $\lambda = 1$, $\varphi(\cdot) = P(x_0, \cdot)$. Of course, other processes also fit into this category.

(6.1) Example

We say that a GSMP has a single set if there exists a state s with which is associated only one clock. In this case, the kernel P has the form

$$P((s,c), A) = p(s'; s, i^*) \prod_{i \in N_{s'}} F(a_i; s', i, s, i^*)$$

so that the right hand side is independent of the clock reading c associated with s. Given that the GSMP passes through $\{s\}$ i.o.,

the chain associated with P is $(A, B, \lambda, \varphi, 1)$ recurrent, where $A = \{(s,c)\}$, $B = S$, $\lambda = 1$, $\varphi(\cdot) = P((s,c), \cdot)$.

An example of a strongly recurrent chain that is not regenerative in the classical sense is given by Example 2.9.

As demonstrated in Section 3, these chains can be embedded in a regenerative environment. Recall that if $\{X_n\}$ is an $(A, B, \lambda, \varphi, 1)$ chain on S, then $\{X_n\}$ can be viewed as a component of the regenerative process $X'_n = (X_n, \delta_n)$ on $S \times \{0,1\}$. By simulating X'_n, one can then obtain estimates for X_n in a straightforward way. We now outline the procedure.

Let f be an S-measurable function with $\pi|f| < +\infty$, and let $r = \pi f$. By simulating X'_n, we obtain a sequence of regeneration times $N_1, N_2, \ldots$. We then form the sequences (with $N_0 = 0$)

$$Y_k = \sum_{j=N_k}^{N_{k+1}-1} f(X_j) \; ; \qquad \bar{Y}_n = \sum_{k=0}^{n} Y_k/(n+1)$$

$$\alpha_k = N_{k+1} - N_k \; ; \qquad \bar{\alpha}_n = \sum_{k=0}^{n} \alpha_k/(n+1)$$

and note that the regenerative character of the X'_n process guarantees that the sequence $\{(Y_k, \alpha_k); k \geq 1\}$ of vectors is i.i.d. The classical regenerative theory then asserts that

$$\hat{r}_n = \bar{Y}_n/\bar{\alpha}_n \to r \qquad P'_\mu, \text{ a.s.} \tag{6.2}$$

This is the well known ratio estimator of regenerative simulation theory. Turning to the question of confidence intervals, let us assume that $\sigma^2 = E(Y_1 - r\alpha_1)^2$ satisfies $0 < \sigma^2 < +\infty$. Then, the i.i.d. structure of the (Y_k, α_k) process gives us the central limit theorem (CLT)

$$P'_\mu\left\{\frac{n^{1/2}(\hat{r}_n - r)}{\sigma/\bar{\alpha}_n} \leq x\right\} \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \tag{6.3}$$

for any initial distribution μ. The value of σ can then be determined by employing standard regenerative estimates (see e.g. Crane and Lemoine [7]). Confidence intervals can then by easily calculated from (6.3).

The limit theorem (6.3) is a result based on simulating a given number of cycles. A second CLT based on number of steps simulated can also be derived. Let

$$U_n = \sum_{k=0}^{n} f(X_k)/(n+1) \quad . \tag{6.4}$$

Then, it can be shown that if $0 < \sigma^2 < +\infty$,

$$P'_\mu\left\{\frac{n^{1/2}(U_n - r)}{\sigma/(E\alpha_1)^{1/2}} \leq x\right\} \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \quad . \tag{6.5}$$

One can of course estimate σ and $E\alpha_1$ using the point estimates indicated above.

Remark

(6.6) For a given strongly recurrent chain, there may exist a multitude of $(A, B, \lambda, \varphi, 1)$ tuples. Consider two such tuples $(A_i, B_i, \lambda_i, \varphi_i, 1)$, $i = 1, 2$. The values of σ and $E\alpha_1$ will in general depend on i. However, (6.5) shows that if $0 < \sigma_1^2, \sigma_2^2 < +\infty$, then

$$\frac{\sigma_1^2}{E_1\alpha_1} = \frac{\sigma_2^2}{E_2\alpha_2} \quad . \tag{6.7}$$

In other words, the ratios of (6.7) are essentially independent of the choice of $(A, B, \lambda, \varphi, 1)$ tuple.

Our next objective is to investigate the choice of $(A, B, \lambda, \varphi, 1)$ tuple from a simulation viewpoint. We suppose, for purposes of the present discussion, that A has been fixed. One now observes that the set B is irrelevant in this context, as a process that is $(A, B, \lambda, \varphi, 1)$ recurrent is clearly $(A, S, \lambda, \hat{\varphi}, 1)$ recurrent, where $\hat{\varphi}$ is the extension of φ given by

$$\hat{\varphi}(E) = \hat{\varphi}(E \cap B) \quad . \tag{6.8}$$

We will not distinguish between $\hat{\varphi}$ and φ in what follows.

Commonly, a simulator, in applying the regenerative method, picks a "regeneration point" with the aim of maximizing the expected number of regenerations over a fixed number of transitions. Now, in our case, the expected number of regenerations over $[0,n]$ is

$$E'_{\mu'}\{\sum_{k=1}^{n} I_{\{X_{k-1} \in A, \delta_k = 1\}}\}/n$$

$$= \sum_{k=1}^{n} P'_{\mu'}\{X_{k-1} \in A,\ \delta_k = 1\}/n$$

$$= \lambda \sum_{k=1}^{n} P'_{\mu'}\{X_{k-1} \in A\}/n \to \lambda\pi(A)$$

so that finding a (λ,φ) pair which maximizes λ and for which the chain is $(A, S, \lambda, \varphi, 1)$ recurrent appears to be a desirable objective.

(6.9) Lemma

Given that a chain is $(A, S, \lambda, \varphi, 1)$ recurrent, there exists a $(\hat{\lambda},\hat{\varphi})$ pair which maximizes $\hat{\lambda}$ over all $(A, S, \lambda, \varphi, 1)$ tuples.

Proof

Let $\hat{\lambda}$ be the λ sup taken over all $(A, S, \lambda, \varphi, 1)$ tuples. Our goal is to show that there exists a measure $\hat{\varphi}$ such that the chain is $(A, S, \hat{\lambda}, \hat{\varphi}, 1)$ recurrent.

Take a sequence of Harris tuples $(A, S, \lambda_n, \varphi_n, 1)$ such that $\lambda_n \nearrow \hat{\lambda}$. Put

$$g_n(x,y) = \lambda_n\left[\frac{d\varphi_n(\cdot)}{dP(x,\cdot)}\right](y)$$

$$h_n(x,y) = \max_{1 \le k \le n} g_k(x,y) \quad .$$

Now, h_n is an increasing sequence of functions so a limit function h exists. Define the measure $\psi_x(\cdot)$ by the formula

$$\psi_x(E) = \int_E h(x,y)\ P(x,\ dy)\ .$$

We shall show that

i) $\psi_x(\cdot)$ is independent of x. Denote the common value by $\hat{\lambda}\hat{\varphi}(\cdot)$.

ii) $(A,\ B,\ \hat{\lambda},\ \hat{\varphi},\ 1)$ is a legitimate Harris tuple.

For i), observe that it is sufficient to show that for each integer n,

$$\int_E h_n(x,y)\ P(x,\ dy) = \int_E h_n(z,y)\ P(z,\ dy)$$

for all $x,\ z \in S$. We prove this fact by induction. For $n = 1$, we have

$$\int_E h_1(x,y)\ P(x,\ dy) = \lambda_1\ \varphi_1(E)$$

which is independent of x. Assume the validity of the induction for $n = k$, and put

$$\psi^k(E) = \int_E h_k(x,y)\ P(x,\ dy)\ .$$

Then, by letting ν be the signed measure $\psi^k(\cdot) - \lambda_{k+1}\ \varphi_{k+1}(\cdot)$, we can apply the Hahn Decomposition Theorem (Royden [21], p. 235) to assert existence of a set F such that F is a positive set (the ν-measure of every subset is non-negative) and F^c is a negative set. Since both $\psi^k(\cdot)$ and $\lambda_{k+1}\ \varphi_{k+1}(\cdot)$ are absolutely continuous w.r.t. each of the measures $P(x,\cdot)$, this implies that

$$h_k(x,y) - g_{k+1}(x,y) \geq 0 , \qquad P(x,\cdot) \text{ a.e. on } F$$

$$h_k(x,y) - g_{k+1}(x,y) \leq 0 \qquad P(x,\cdot) \text{ a.e. on } F^c$$

Now, $h_{k+1}(x,y) = \max\{h_k(x,y), g_{k+1}(x,y)\}$ so that for each x, we have

$$\int_E h_{k+1}(x,y)\, P(x, dy)$$

$$= \int_{E\cap F} h_k(x,y)\, P(x, dy) + \int_{E\cap F^c} g_{k+1}(x,y)\, P(x, dy)$$

$$= \psi^k(E \cap F) + \lambda_{k+1}\, \varphi_{k+1}(E \cap F^c)$$

which is independent of x, proving i). We now turn to ii). We first note that by Monotone Convergence,

$$\hat{\lambda}\hat{\varphi}(S) = \int_S h(x,y)\, P(x, dy)$$

$$= \lim_{h \to \infty} \int_S h_n(x,y)\, P(x, dy)$$

$$\geq \lim_{n \to \infty} \max_{1 \leq k \leq n} \int_S g_k(x,y)\, P(x, dy)$$

$$= \lim_{n \to \infty} \lambda_n = \hat{\lambda}$$

so that $\hat{\varphi}(S) \geq 1$. We now observe that by Remark 3.3, $0 \leq h(x,y) \leq 1$ $P(x,\cdot)$ a.e. so

$$\hat{\hat{\lambda}}\hat{\varphi}(E) = \int_E h(x,y)\ P(x,dy) \leq P(x,E)\ .$$

Finally, putting $\varphi'(\cdot) = \hat{\varphi}(\cdot)/\hat{\varphi}(S)$, $\lambda' = \hat{\hat{\lambda}}\hat{\varphi}(S)$, we see that $(A, S, \lambda', \varphi', 1)$ is a legitimate Harris tuple. But $\hat{\lambda}$ was the λ sup, and so $\lambda' = \hat{\lambda}$ and hence $\hat{\varphi}(S) = 1$, forcing $\varphi' = \hat{\varphi}$.

We shall now discuss a constructive means of determining $\hat{\lambda}$ and $\hat{\varphi}$, under certain restrictions. Suppose that S is a metric space, and that A contains a countable subset C which is dense in A. Assume also that the kernel P has the form

$$P(x,E) = \sum_{z \in I_x} a_{z,x}\ \delta_z(E) + \int_E p(x,y)\ \mu(dy)$$

where I_x is a countable set, $p(\cdot,y)$ is continuous for each y, and $\mu(\{z\}) = 0$ for each $z \in S$.

To share that a pair $(\hat{\lambda}, \hat{\varphi})$ is optimal, it is sufficient to prove that the measure $\psi(\cdot) = \hat{\hat{\lambda}}\hat{\varphi}(\cdot)$ dominates all other finite measures τ that satisfy $\tau(\cdot) \leq P(x,\cdot)$ for every x in A. Our candidate for ψ is

$$\psi(E) = \sum_{z \in I} a_z\ \delta_z(E) + \int_E p(y)\ \mu(dy) \tag{6.10}$$

where $I = \bigcap_{x \in A} I_x$, $a_z = \inf_{x \in A} a_{z,x}$, $p(y) = \inf_{x \in A} p(x,y)$. We observe that ψ is well-defined since our restriction on A and continuity assumption on $p(\cdot,y)$ imply that

$$p(y) = \inf_{x \in A} p(x,y) = \inf_{x \in C} p(x,y)$$

so that $p(\cdot)$, being the infinum of countably many functions, is measurable.

To show that ψ is a feasible solution, it is necessary to prove that $\psi(\cdot) \leq P(x,\cdot)$ for each x in A. But this follows as a consequence of the inequality

$$\int_E p(y)\ \mu(dy) = \int_E \inf_{x \in A} p(x,y)\ \mu(dy)$$

$$\leq \inf_{x \in A} \int_E p(x,y)\ \mu(dy) \quad .$$

As for the optimality, let τ be any other finite measure satisfying $\tau(\cdot) \leq P(x,\cdot)$. Then, it is easily checked that

$$\tau_1(E) = \tau(E \cap I) \leq \sum_{z \in I} a_{z,x}\ \delta_z(E) \ , \qquad \forall\ x \in A$$

(6.11)

$$\tau_2(E) = \tau(E \cap I^c) \leq \int_E p(x,y)\ \mu(dy) \ , \qquad \forall\ x \in A \ .$$

For τ_1, this clearly implies that

$$\tau_1(E) \leq \sum_{z \in I} a_z\ \delta_z(E) \quad . \tag{6.12}$$

As for τ_2, we have that $\tau_2 << \mu$ so, letting $\tau_2(y)$ be the derivative of τ_2 w.r.t. μ, we have

$$\tau_2(y) \leq p(x,y) \qquad \mu \text{ a.e.}$$

for each $x \in A$. But this means that

$$W = \{y:\tau_2(y) \leq p(x,y) \ \forall\ x \in C\}$$

$$= \{y:\tau_2(y) \leq \inf_{x \in C} p(x,y)\}$$

has $\mu(W^c) = 0$. Hence, $\tau_2(y) \leq p(y)$ μa.e. This, together with (6.11) and (6.12), implies that $\tau(E) \leq \psi(E)$, and so ψ is optimal.

To recover $\hat{\lambda}$, $\hat{\varphi}$, put $\hat{\lambda} = \psi(S)$, $\hat{\varphi}(\cdot) = \psi(\cdot)/\hat{\lambda}$.

Remarks

(6.13) The above discussion kept A fixed. If a simulator has several candidates A_i for A, then (at least theoretically) one could run trial simulations to estimate $\pi(A_i)$, and determine the optimal $\hat{\lambda}_i$ for each of the A_i. One could then choose the A_i maximizing $\hat{\lambda}_i \pi(A_i)$ as the "best" A.

(6.14) One could further increase the frequency of regeneration by employing the followng method. Let $q_x(y)$ be the derivative of $\hat{\varphi}$ w.r.t. $P(x,\cdot)$, and define

$$\frac{1}{\lambda_x} = \|q_x(\cdot)\|$$

so that $P(x,E) \geq \lambda_x \hat{\varphi}(E)$. Observe that $\lambda_x \geq \hat{\lambda}$ for every x in A, and that the decomposition

$$P(x,E) = \lambda_x \hat{\varphi}(E) + (1-\lambda_x)\, Q(x,E)$$

is still possible. One could now follow precisely the method of (3.9) to (3.10) to embed $\{X_n\}$ in a regenerative M.C. X'_n. The asymptotic frequency of regeneration for this chain would be

$$\int_S \lambda(x)\, \pi(dx)$$

which is clearly greater than $\hat{\lambda}\pi(A)$.

Example 2.8 continued

Consider the M/M/1 version of the $\{W_n\}$ chain. Here, the density f is given by

$$f(y) = \begin{cases} \frac{\alpha}{(\alpha+1)}\, e^{\alpha y}, & y < 0 \\ \frac{\alpha}{(\alpha+1)}\, e^{-y}, & y \geq 0 \end{cases}$$

where $0 < \alpha < 1$. Consider a recurrence interval of the form $A = [0,b]$, $b > 0$. All of the criterion of our optimality discussion are satisfied so, putting $\psi = \hat{\lambda}\hat{\varphi}$,

$$\psi(E) = \delta_0(E)\ F(-b) + \int_E p(y)\ dy$$

where

$$p(y) = \inf_{x\ \in\ [0,b]} \{f(y-x)\}$$

$$= \inf_{u\ \in\ [y-b,y]} \{f(y)\}$$

$$= \min\{f(y-b),\ f(y)\}$$

$$= \begin{cases} \frac{\alpha}{(\alpha+1)}\ e^{\alpha(y-b)}\ , & 0 \leq y \leq \frac{\alpha b}{\alpha+1} \\ \\ \frac{\alpha}{(\alpha+1)}\ e^{-y}\ , & y > \frac{\alpha b}{\alpha+1}\ . \end{cases}$$

Of course, $\hat{\lambda} = F(-b) + \int_0^\infty p(y)\ dy$, and $\hat{\varphi}(\cdot) = \psi(\cdot)/\hat{\lambda}$.

The measures $Q(x,\cdot)$ can also be calculated. Defining H, h through

$$(1-\hat{\lambda})\ Q(x,E) = H(x)\ \delta_0(E) + \int_E h(x,y)\ dy$$

we see that $H(x) = F(-x) - F(-b)$, $h(x,y) = f(y-x) - p(y)$. For $0 \leq x \leq \lambda b/(\lambda+1)$, $h(x,y)$ takes the form

$$h(x,y) = \begin{cases} \frac{\alpha e^{\alpha y}}{(\alpha+1)} \left[e^{-\alpha x} - e^{-\alpha b}\right], & 0 \le y \le x \\ \frac{\alpha}{\alpha+1} \left[e^{-(y-x)} - e^{\alpha(y-b)}\right], & x \le y \le \frac{\lambda b}{\lambda+1} \\ \frac{\alpha e^{-y}}{\alpha+1} \left[e^{x} - 1\right], & y \ge \frac{\lambda b}{\lambda+1} \end{cases}$$

The function $h(x,y)$ takes on an equally complicated form over each of the intervals $[\lambda b/(\lambda+1), b]$, $[b, +\infty)$.

The above example indicates that the computation involved in calculating $Q(x,\cdot)$ may, in general, be formidable. However, from a simulation stand point, explicit knowledge of $Q(x,\cdot)$ is unnecessary. One can always obtain deviates distributed according to $Q(x,\cdot)$ by utilizing the following acceptance-rejection method.

We assume that the assumptions on S, A, and P required to obtain (6.10) are in force. The first step is to obtain a deviate Z distributed according to $P(x,\cdot)$. One then puts

$$\bar{p}(Z) = \begin{cases} a_z/a_{z,x}, & \text{if } Z = z \in I_x \\ p(Z)/p(x,Z), & \text{otherwise} \end{cases} \tag{6.15}$$

and generates a Bernoulli random variable I such that $P\{I = 1|Z\} = \bar{p}(Z) = 1 - \bar{q}(Z)$. We claim that

$$P\{Z \in E | I = 1\} = \hat{\varphi}(E)$$

$$P\{Z \in E | I = 0\} = Q(x,E) \quad .$$

To see this, observe that

$$P\{Z \in E,\ I = 1\}$$

$$= P\{Z \in E \cap I_x,\ I = 1\} + P\{Z \in E \cap I_x^c,\ I = 1\}$$

$$= \sum_{z \in E \cap I_x} a_{z,x}\, P\{I = 1 | Z = z\} + \int_E P\{I = 1 | Z = z\}\, p(x,z)\, \mu(dz)$$

$$= \sum_{z \in E \cap I_x} a_z + \int_E p(z)\, \mu(dz) = \psi(E)$$

so that $P\{I = 1\} = P\{Z \in S,\ I = 1\} = \psi(S) = \hat{\lambda}$. Hence, $P\{Z \in E | I = 1\} = \hat{\varphi}(E)$. For $Q(x,\cdot)$, note that $P\{Z \in E,\ I = 0\} = P\{Z \in E\} - P\{Z \in E,\ I = 1\}$ $= P(x,E) - \psi(E) = Q(x,E)\ (1-\hat{\lambda})$, and that $P\{I = 0\} = 1-\hat{\lambda}$.

Our final comments in this section focus on implementation of the above method in an actual simulation. The method requires a simulator to determine a Harris tuple $(A, B, \lambda, \varphi, 1)$. The simulator's intuition for the model to be simulated would be a valuable tool in obtaining appropriate choices for the set A. One would have to consider whether the set A is hit infinitely often, as well as whether the kernel could be expected to be reasonably smooth over A.

The determination of λ and φ is a more exacting matter. While the mathematical conditions involved in applying (6.10) to finding λ and φ are fairly mild (from a practical standpoint), such an application does require an explicit knowledge of the kernel P, over the set A. However, most simulations of complex stochastic systems are generated through an event scheduling approach that does not require an explicit form for the transition probabilities (see e.g., Crane and Lemoine [7], p. 72-83). Therefore, the simulator will generally be faced with the task of determining the structure of the kernel P, a task which could potentially be quite time consuming. However, it is to be emphasized that an explicit form for P is required only over the set A.

Once λ and φ have been found, simulating the X_n' chain would require only a moderate additional amount of work. To obtain X_n', one would generate a sample path of the process X_n in the normal way. At each transition of X_n, one would also generate a Bernoulli random variable δ_n with parameter $\bar{p}(X_n)$. The value of $\bar{p}(x)$ is given by (6.15) over the set A, while its value over A^c is λ identically. The pair (X_n, δ_n) would then constitute the process X_n'.

7. Weakly Recurrent Harris Chains

We shall say that a Harris chain is weakly recurrent if it is not strongly recurrent. In this section, we shall investigate a means for embedding a weakly recurrent chain in a process that can be statistically analyzed in a more manageable way.

In Section 3, we gave an example (Example 3.17) of a process which is weakly recurrent. We now give a less artificial example. Consider the chain associated with a GSMP that contains no single sets, and suppose that the process is $(B_1, B_2, \alpha, \psi, 1)$ recurrent. For each (s,c) in the state space S of the chain, the kernel $P((s,c), \cdot)$ contains an indicator function. Then, since $\psi << P((s,c), \cdot)$ for some (s,c), ψ must be supported on that part of S for which the clock $i \in O_s$, is set at the value c_i^* (see (2.11)). Clearly, if ψ is supported on such a set, the chain must pass through the set infinitely often. But, in general, a clock will never be set stochastically at the same value twice and hence the chain cannot visit such a set infinitely often. Hence, the chain associated with a GSMP without single set will generally be only weakly recurrent. This motivates our present discussion.

As discussed in Section 3, the randomization $\hat{P}$ of a weakly recurrent kernel P with invariant probability π is strongly recurrent. If one were interested in an estimate of πf for the kernel P, one could then simulate the chain associated with $\hat{P}$ using the techniques of Section 6 for strongly recurrent chains to obtain a

confidence interval. The most natural way to simulate the randomized chain would be to employ its interpretation as a random sampling of the original chain (see Remark 3.19). In other words, one would simulate the original chain $\{X_n\}$, as well as an independent i.i.d. sequence $\{Z_n\}$ of geometric random variables with parameter p. Then, by putting $Y_n = n + Z_1 + Z_2 + \cdots + Z_n$, the process $\{X_{Y_n}\}$ would have exactly the distribution of the randomized process. The difficulty is that one only retains a proportion p of the $\{X_n\}$ sample path in forming confidence intervals based on the $\{X_{Y_n}\}$ chain. Since this is clearly undesirable, we examine a second method of simulating weakly recurrent chains.

We shall show that under mild assumptions on $(S,\mathcal{S})$, a weakly recurrent chain $\{X_n\}$ can be embedded in a process that is loosely regenerative in the sense of Smith [22]. We shall be more precise about this later.

Let the weakly recurrent chain $\{X_n\}$ be $(A, \mathcal{A}, \lambda, \varphi, n)$ recurrent. We start by considering the process $Z_m = (X_{nm+1}, \ldots, X_{n(m+1)})$ which lives on the state space S^n, equipped with product σ-field $\mathcal{S}^n$, and transition kernel K given by

$$(7.1) \qquad K((x_1, \ldots, x_n), E_1 \times \cdots \times E_n)$$

$$= \int_{E_1} P(x_n, dy_1) \int_{E_2} P(y, dy_2) \cdots \int_{E_n} P(y_{n-1}, dy_n) \quad .$$

Then, for an initial probability σ on $(S^n, \mathcal{S}^n)$, we can construct a probability tuple $(\Omega^n, F^n, K_\sigma)$ for which the random variables Z_m are obtained by projection (see Section 2).

The original chain $\{X_n\}$, which has initial probability μ (say), is recovered by setting

$$(7.2) \quad \bar{\mu}(E_1 \times \cdots \times E_n) = \int_S \mu(dx) \int_{E_1} P(x,dy_1) \cdots \int_{E_n} P(y_{n-1}, dy_n)$$

The process $\{Z_m; m \geq 0\}$ governed by $K_{\bar{\mu}}$ then corresponds precisely to $\{X_n\}$ under P_μ. Now, it can be shown (Athreya and Ney [1]) that if $\{X_n\}$ is initiated with probability μ concentrated on A, then the process X_{nk} visits A infinitely often. This clearly implies that the process Z_m visits $S^{m-1} \times A$ infinitely often under $K_{\bar{\mu}}$.

We assume for the remainder of this section that S is a complete, separable metric space, with Borel σ-field $\mathcal{S}$. Now, we have that

$$P^n(x,E) \geq \lambda\varphi(E)$$

for every x in A and E in $\mathcal{A}$. Again, we can perform the decomposition

$$P^n(x,E) = \lambda\varphi(E) + (1-\lambda)\, Q(x,E)$$

where Q is a transition kernel on $(S, \mathcal{S})$. We want to perform a similar decompositon on the kernel K. Now, since S is a complete, separable metric space, $(S^n, \mathcal{S}^n)$ is a Borel space (see Breiman [5], p. 79, 402) and so a regular conditional probability

(7.3) $$P_x(X_1 \in E_1, \ldots, X_{n-1} \in E_{n-1} | X_n = y)$$

exists for every y in S. That means that for each $y \in S$, (7.3) is a probability on $(S^{n-1}, \mathcal{S}^{n-1})$, whereas for each tuple $(E_1, \ldots, E_{n-1})$, (7.3) is a $\mathcal{S}$-measurable version of the conditional probability. Define the family of measures $\bar{\varphi}(x, \cdot)$, $\bar{Q}(x, \cdot)$ by the formulas

(7.4) $$\bar{\varphi}(x, E_1 \times E_2 \times \cdots \times E_n)$$

$$= \int_{E_n} P_x(X_1 \in E_1, \ldots, X_{n-1} \in E_{n-1} | X_n = y)\, \varphi(dy)$$

$$\bar{Q}(x, E_1 \times E_2 \times \cdots \times E_n)$$

$$= \int_{E_n} P_x(X_1 \in E_1, \ldots, X_{n-1} \in E_{n-1} | X_n = y)\, Q(x, dy)$$

for each x in A. Observe that for each $(x_1, \ldots, x_n) \in S^{n-1} \times A$, we have

(7.5) $$K((x_1, \ldots, x_n), E_1 \times \cdots \times E_n)$$

$$= \lambda \bar{\varphi}(x_n, E_1 \times \cdots \times E_n) + (1-\lambda)\, \bar{Q}(x_n, E_1 \times \cdots \times E_n) .$$

We can now describe the embedding for Z_m. Let $S' = S^n \times \{0,1\}$ and let $\mathcal{S}'$ be the associated product σ-field. Define a kernel K' relative to $(S', \mathcal{S}')$ by

$$K'(((x_1, \ldots, x_n), \delta), E_1 \times \cdots \times E_n \times \{0\})$$

$$= \begin{cases} (1-\lambda)\, K((x_1, \ldots, x_n), E_1 \times \cdots \times E_n), & (x_1,\ldots,x_n) \notin S^{n-1} \times A \\ (1-\lambda)\, \bar{Q}(x_n, E_1 \times E_2 \times \cdots \times E_n), & (x_1,\ldots,x_n) \in S^{n-1} \times A \end{cases}$$

$$K'(((x_1, \ldots, x_n), \delta), E_1 \times \cdots \times E_n \times \{1\})$$

$$= \begin{cases} \lambda K((x_1, \ldots, x_n), E_1 \times \cdots \times E_n), & (x_1,\ldots,x_n) \notin S^{n-1} \times A \\ \lambda\bar{\varphi}(x_n, E_1 \times \cdots \times E_n), & (x_1,\ldots,x_n) \in S^{n-1} \times A \end{cases}$$

We extend measures σ on $(S^n, \mathcal{S}^n)$ to measures σ' on $(S', \mathcal{S}')$ as in (3.4). Then, by letting $Z'_m = (Z_m, \delta_m)$ be the canonical M.C. corresponding to S' and K', we see that

i) for any σ on $(S^n, \mathcal{S}^n)$, the marginal distribution (w.r.t. K'_σ,) of the coordinate process Z_m of the M.C. Z'_m is the same as that of the original chain Z_m (w.r.t. K_σ).

ii) the random variables $\{\delta_n\}$ constitute an i.i.d. Bernoulli sequence with parameter λ.

Define $N_1 = \inf\{m \geq 1 : Z_{m-1} \in S^{n-1} \times A, \delta_m = 1\}$ and put $N_k = \inf\{m > N_{k-1} : Z_{m-1} \in S^{n-1} \times A, \delta_m = 1\}$. We first note that there are an infinite number of N_k, provided $\sigma = \bar{\mu}$, where μ is concentrated

on A. Since our discussion is motivated by ergodic properties of $\{X_n\}$, it is no restriction to only consider such μ, for by (5.9) the initial position washes out in the limit.

We now consider the sequence of random vectors

$$\hat{Z}_k = (Z'_{N_k+1}, \ldots, Z'_{N_{k+1}}) , \qquad k \geq 1 .$$

Now, the sequence $\{\hat{Z}_k;\ k \geq 1\}$ is identically distributed. By the Strong Markov property, this will follow if the sequence consisting of the first component of Z_{N_k+1} is identically distributed. The distribution of this component is given by

$$K'_{\mu}, \{X_{n(N_k+1)+1} \in E\} = \int_S P(y,E)\ \varphi(dy) \tag{7.6}$$

proving our claim. However, the sequence of vectors $\hat{Z}_k$ is not independent as the distribution of Z_{N_k+1} clearly depends on that of Z_{N_k}. It is true, though, that the sequence $\{\hat{Z}_k;\ k \geq 1\}$ is 1-dependent. In other words, $\{\hat{Z}_1, \ldots, \hat{Z}_{k-1}\}$ is independent of $\{\hat{Z}_{k+1}, \ldots, \hat{Z}_{k+m}\}$ for $m \geq 1$. This is basically a consequence of the independence of Z_{N_k-1} and Z_{N_k+1}, which we now endeavor to prove. Evidently

$$K'_{\mu}, \{Z_{N_k+1} \in E | Z_{N_k-1}\}$$

$$= E'_{\mu}, \{K'_{\mu}, \{Z_{N_k+1} \in E | F'_{N_k}\} | Z_{N_k-1}\} ,$$

where F'_{N_k} is the σ-field generated by the stopping time N_k along the Z' sample path. By the Strong Markov property, the above is just

$$E'_{\underline{\mu}}, \{K'_{Z'_{N_k}} \{Z_1 \in E\} | Z_{N_k-1}\} \quad .$$

But observe that the interior probability depends on Z'_{N_k} only through $X_{n(N_k+1)}$ and this later random variable always has distribution $\varphi(\cdot)$, completing the argument.

Remark

(7.7) The existence of a sequence of stopping times N_k for which Z_{N_k-1} and Z_{N_k+1} are independent guarantees that $\{Z_m\}$ is loosely regenerative in the sense of Smith [22]. To incorporate such a process into his framework, consider the Z_m process at each of the stopping times $N'_k = N_k+1$. Then, the history of the process up to time N'_k-2 has no value in predicting the future evolution of Z_m, and so this process is loosely regenerative.

We will now obtain a CLT for the Z'_n process. Let f be S-measurable with $\pi|f| < +\infty$, where π is a probability assumed invariant for $\{X_n\}$. Extend f to a function $\bar{f}$ on S^n by the formula

$$\bar{f}(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} f(x_i) \quad .$$

Then, putting $r = \pi f$, we can define the sequences

$$Y_k = \sum_{j=N_k+1}^{N_{k+1}} \bar{f}(Z_j)$$

$$\alpha_k = N_{k+1} - N_k$$

$$V_k = Y_k - r\alpha_k$$

which we note are all 1-dependent. Also, since the sequence $\{\hat{Z}_k\}$ is identically distributed, all three of the above sequences are strictly stationary. Furthermore, the sequence

$$\sum_1^m Y_k / \sum_1^m \alpha_k$$

is merely the sequence $\{\sum_1^{\ell} f(X_i)/\ell\}$ stopped at the random times nN_{m+1} Hence by the nature of a.s. convergence, it must be that

$$\hat{r}_m = \sum_1^m Y_k / \sum_1^m \alpha_k \to r \qquad K'_{\bar{\mu}}, \text{ a.s.} \tag{7.8}$$

It should be, then, that $E'_{\bar{\mu}} V_k = 0$. Thus, given that $\bar{\sigma}^2 = E'_{\bar{\mu}} V_k^2$ satisfies $0 < \bar{\sigma}^2 < +\infty$, we can apply Theorem 20.1 of Billingsley [3] for φ-mixing processes to obtain the C.L.T.

$$K'_{\bar{\mu}}, \left\{ \sum_{k=1}^m V_k \Big/ \sqrt{m}\sigma \le x \right\} \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

where $\sigma^2 = \underline{E}'_{\underline{\mu}}, V_k^2 + 2\underline{E}'_{\underline{\mu}}, V_k V_{k+1}$. Consequently

$$n^{1/2} \frac{[\hat{r}_n - r]}{(\sigma/\bar{\alpha})} \tag{7.9}$$

also satisfies a C.L.T., where $\bar{\alpha}$ is the sample average of $\{\alpha_k\}$. To apply (7.9) to confidence interval estimation, it would be necessary to estimate σ^2. Expanding σ^2 as

$$\underline{E}'_{\underline{\mu}}, \{Y_k^2 - 2r\alpha_k + r^2\alpha_k^2 + Y_k Y_{k+1} - r\alpha_k Y_{k+1} - r\alpha_{k+1} Y_k + r^2\alpha_k\alpha_{k+1}\} \tag{7.10}$$

we see that each term can be estimated in a fairly straightforward way. Results (7.8) through (7.10) then form the basis for a confidence interval estimation technique based on the Z'_m process.

The above argument, while somewhat loose, points to a simulation technique for estimation of "steady-state" values πf. We now briefly discuss implementation of the method.

A first step would be to determine the (λ,φ) pair corresponding to the tuple $(A, A, \lambda, \varphi, n)$. Given that an explicit representation of P^n was available over the set A, one could find (λ,φ) through formula (6.10), provided that the assumptions on A and P^n required for validity of (6.10) were in force. While we have already mentioned in Section 6 the difficulties in finding P explicitly, the problems to be surmounted in doing the same for P^n would surely be even more formidable. This is because the simulator would have to follow the process through n transitions in order to find P^n, and this could conceivably be a very complex task.

We now turn to generation of sample paths for the process Z_n', given knowledge of a (λ,φ) pair. The simulation would begin at a point picked from the set A. The simulator would then generate a sample path for the original process $\{X_m\}$ in the usual way. However, were X_{nk} to be in A, one would record its position x in that set. The simulation would then continue to run in the normal fashion until $X_{n(k+1)}$ was generated. The acceptance-rejection method of (6.15) would then be employed in order to determine whether $X_{n(k+1)}$ came from φ or $Q(x,\cdot)$. Given that $X_{n(k+1)}$ was distributed according to φ, N_1 would be set to k. The above chain of events would then repeat itself to generate the complete sequence $N_2, N_3, \ldots$.

The final step for the simulator would then be to use (7.8) through (7.10) to form a confidence interval for the quantity πf of interest.

Appendix

Our objective here is to show that

$$\hat{P}_x\{T_A < +\infty\} = 1$$

for every x in S, where $\hat{P}$ is the randomized version of a Harris kernel P (with invariant probability π) and $A \in S$ is a set such that

$$P_x\{T_A < +\infty\} = 1$$

for every x in S. We start with a couple of definitions. We say that $\Lambda \in F$ is an invariant event if $\theta^{-1}(\Lambda) = \Lambda$, and denote the σ-field of invariant events by I. If Z is I-measurable, one calls Z an invariant random variable. Finally, a S-measurable function h is said to be P-harmonic if

$$h(x) = \int_S h(y)\, P(x, dy)$$

for every x in S.

(A.1) Lemma

Let Z be a bounded invariant random variable. Then, $h(x) = \hat{E}_x Z$ is P-harmonic.

Proof.

Since Z is invariant, $Z \circ \theta_j = Z$ for all j, and so

$$\hat{E}_x Z = \hat{E}_x (Z \circ \theta_j)$$

$$= \hat{E}_x \{\hat{E}_{X_j} Z\}$$

$$= \int_S \hat{E}_y Z \cdot \hat{P}^j(x, dy)$$

by the Markov property. Now, it is easy to calculate that

$$\hat{E}_x Z = \sum_{n=0}^{\infty} \binom{j+n-1}{n} p^j (1-p)^n \int_S \hat{E}_y Z \cdot P^{n+j}(x, dy) \quad . \tag{A.2}$$

Put

$$a_i = \int_S \hat{E}_y Z \cdot P^i(x, dy)$$

$$f(\lambda) = a_1 + a_2\lambda + a_3\lambda^2 + \cdots \quad . \tag{A.3}$$

Clearly, since $|a_i| \leq \|Z\|$ for all i, the radius of convergence of f is at least one, and hence f is analytic on any interval $[-t,t]$ for $t < 1$. Hence, for any $\lambda \in (-1, 1)$, we have for $j \geq 1$

$$f^{(j)}(\lambda) = \sum_{n=0}^{\infty} \frac{(n+j)!}{n!} a_{n+j+1} \lambda^n$$

so that for $p > 0$ (putting $q = 1-p$)

$$f^{(j)}(q) = \frac{j!}{p^{j+1}} \sum_{n=0}^{\infty} \binom{j+n}{n} p^{j+1} q^n a_{n+j+1}$$

$$= \frac{j!}{p^{j+1}} \hat{E}_x Z \qquad \text{(by (A.1))}$$

$$= \frac{j!}{p^j} f(q) \qquad \text{(by (A.2), (A.3))} .$$

But, since f is analytic at q,

$$f(q+h) = \sum_{j=0}^{\infty} f^{(j)}(q) \frac{h^j}{j!} = f(q) \left(\frac{p}{p-h}\right)$$

which is valid at least for $|h| < p$. Thus

$$f(\lambda) = f(q) \left(\frac{p}{1-\lambda}\right) \tag{A.4}$$

is valid for $q-p < \lambda < 1$. But f is analytic on any subinterval of $(-1, 1)$, so the representation (A.4) is valid at $\lambda = 0$. So,

$$f(0) = pf(q)$$

i.e.,

$$\int_S \hat{E}_y Z \cdot P(x, dy) = \hat{E}_x Z \quad . \qquad \|$$

Now, it can be shown that every bounded $\hat{P}$-harmonic function h has the form $h(x) = \hat{E}_x Z$ for some $\mathcal{I}$-measurable Z (Revuz [20], p. 82). Hence, Lemma A.1 can be reformulated as:

Every bounded $\hat{P}$-harmonic function h is P-harmonic .

(A.5) Lemma

The bounded P-harmonic functions are constant.

Proof.

Let h be a bounded P-harmonic function. Then, by (5.3)

$$h(x) = \frac{1}{n+1} \sum_{k=0}^{n} (P^k f)(x) \to \int_S h(y)\, \pi(dy)$$

so that h is identically constant. ‖

Hence, the bounded $\hat{P}$-harmonic functions are constant. This is known to imply that every set is either recurrent or transient for $\hat{P}$ (Revuz [20], p. 83). A set A is said to be recurrent (transient) for $\hat{P}$ if $\hat{P}_x\{X_n \in A \text{ i.o.}\} = 1$ $(= 0)$ for every $x \in S$.

Now, existence of an invariant probability π for P has been assumed. This measure π is then trivially invariant for $\hat{P}$, and so θ is a measure preserving transformation on the probability space $(\Omega, F, \hat{P}_\pi)$. Then, a well-known ergodic result (see e.g. Lamperti [15], p. 91) states that

$$\hat{P}_\pi\{X_n \in A \text{ i.o.}\} \geq \pi(A) > 0$$

implying that A must be a recurrent set under $\hat{P}$.

Remark

(A.6) Note that existence of an invariant π for $\hat{P}$ is sufficient for the above argument to be valid, for then

$$\pi(A) = \int_S \hat{P}(x,A)\, \pi(dx) \quad .$$

We now observe that since $\{X_n\}$ hits A P_x a.s., it must be that $\hat{P}(x,A)$ is positive for every x in S, and consequently $\pi(A)$ is positive. This is the only modification required.

Acknowledgement

The author wishes to thank Professor Donald L. Iglehart for his valuable suggestions during this research. Conversations with Laurent Cantaluppi also were of substantial benefit. The author would also like to thank Gail Lemmond Stein for her excellent typing of the manuscipt.

References

1. Athreya, K.B. and P. Ney (1977). A new approach to the limit theory of Markov chains. Preprint.

2. Athreya, K.B. and P. Ney (1978). A new approach to the limit theory of Markov chains. Trans. Amer. Math. Soc. 245, 493-501

3. Billingsley, P. (1968). Convergence of Probability Measures. John Wiley and Sons, New York.

4. Breiman, L. (1960). A strong law of large numbers for a class of Markov chains. Ann. Math. Statist. 31, 801-803.

5. Breiman, L. (1968). Probability. Addison-Wesley, Reading, Mass.

6. Copson, E.T. (1968). Metric Spaces. Cambridge University Press, Cambridge.

7. Crane, M.A. and A.J. Lemoine. (1978). An Introduction to the Regenerative Method for Simulation. Springer-Verlag, Berlin.

8. Dellacherie, C. and P.A. Meyer (1978). Probabilities and Potential. North-Holland Publishing Company, New York.

9. Doob, J.L. (1953). Stochastic Processes. John Wiley and Sons, New York.

10. Feller, W. (1971). An Introduction to Probability, Vol. 2. John Wiley and Sons, New York.

11. Fossett, L.D. (1979). Simulating generalized semi-Markov processes. Technical Report 47, Department of Operations Research, Stanford University.

12. Glynn, P.W. (1980). Simulation of Harris recurrent Markov chains. Forthcoming Technical Report, Department of Operations Research, Stanford University.

13. Harris, T.E. (1956). The existence of stationary measures for certain Markov processes. Proc. Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. 2. Jerzey Neyman, ed., University of California Press, Berkeley.

14. Karr, A.F. (1975). Weak convergence of sequence of Markov chains. Z. Wahrscheinlichkeitstheorie ver Geb. 33, 41-48.

15. Lamperti, J. (1977). Stochastic Processes. Springer-Verlag, Berlin.

16. Neveu, J. (1965). Mathematical Foundations of the Calculus of Probability, Holden-Day, San Francisco.

17. Nummelin, E. (1978). A splitting technique for Harris recurrent Markov chains. Z. Wahrscheinlichkeitstheorie ver Geb. 43. 309-318.

18. Orey, S. (1959). Recurrent Markov chains. Pacif. J. Math. 9, 805-827.

19. Orey, S. (1971). Limit Theorems for Markov Chain Transition Probabilities, Van Nostrand, New York.

20. Revuz, D. (1975). Markov Chains, North-Holland, Amsterdam.

21. Royden, H.R. (1968). Real Analysis. MacMillian, New York.

22. Smith, W.L. (1955). Regenerative stochastic processes. Proc. Roy. Soc. London Ser. A. 232, 6-31.

23. Whitt, W. (1980). Continuity of generalized semi-Markov processes. To appear in Math. Oper. Res.